Critical Thinking Instruction Through Project-Based Learning in Chinese EFL Classes- A Case Study in HE

Submitted for the Degree of Doctor of PhilosophyAt the University of Northampton

2022

Qiuling Bi

ACKNOWLEDGMENTS

I am grateful to many people, who have supported and encouraged me throughout my PhD journey, especially during the Covid-19 period.

I especially wish to express my sincere appreciation to my supervision team first. Dr. Qian Zhang, as my first supervisor, you are the only one who accompanied me from the very beginning until the end of my research. Her patience, professionalism and expertise were a constant inspiration. Without her, it would be harder to complete this study.

I am full of gratitude to the team director and my current second supervisor Dr. Emel Thomas. Since joining this team, her committed leadership, continued encouragement and constructive feedback made me work more efficiently. Without her, it would be longer to complete this research.

Thank you to my former team director, Dr. Anna Cox and second supervisor, Dr. Amanda O'Shea. Their valuable suggestions always encouraged me to persist my research.

Thank you to Dr. Cristina Devecch and Simone Apel. Their guidance and valuable suggestions about doing presentations helped me to gain more academic public speaking experiences.

Special thanks in particular to my research subjects, who volunteered willingly and enthusiastically.

My husband, Tao Wang, who always loves me, worked very hard to support me in pursuing my dream and tolerated my absences from normal family life during the four years.

Very special thanks to my son, who never complained and always supported me to reduce my sense of anxiety being absent in his valuable middle school life.

My parents always supported me and helped my husband take care of my son, reducing his burden and making him concentrate on work. Their concern, encouragement and love always inspired me to engage in my study.

I could not have done this project without the support of my friends, in particular, Ming Nie, Cuixia Han, Ge Yan, Lin Yuan, Li Li, Yue Yang, Han Lu, Tianyang Zhu, Xianrong Wang and Kai Wang. Thank you for your listening, encouragement and support.

ABSTRACT

Critical thinking is a fundamental skill for undergraduates in the twenty first century. The Chinese Ministry of Education has conducted a series of national curriculum reforms since 2017, aimed to cultivate university students' higher-order thinking skills. Although research in critical thinking pedagogy has attracted increasing attention in recent years, there is no consensus on it (Nagjie *et al*., 2020; Zou and Lee, 2021). This case study intended to explore whether CT-oriented PBL pedagogy is applicable and practical to developing students' thinking using a mixed method.

The setting was undertaken in a Chinese medical university with a group of 98 first-year undergraduates who studied College English as one of their modules. A 15-week CT oriented PBL teaching intervention programme was designed and delivered to the participants. Both quantitative and qualitative data was collected before, during and after the PBL intervention using the California Critical Thinking measurements (the simplified Chinese version), students' writing projects, questionnaires and interviews.

Three findings of this study indicate that CT-oriented PBL intervention is applicable and effective for students' thinking enhancement. Firstly, the results of CT tests demonstrated that students significantly developed generic CT skills (i.e. Evaluation, Analysis and Inference) and dispositions (i.e. Inquisitiveness, Confidence in Reasoning, Open-mindedness and Analyticity) after PBL intervention. Data from their writing projects during and after PBL instruction confirmed these findings of a significant increase in the use of content-specific thinking skills. In addition, the result from interviews with learners after PBL were partly in line with the finding in developing thinking skills. Secondly, results from the post-PBL questionnaires suggested that students held positive attitudes towards the effectiveness of CT strategies (i.e. thinking maps, group discussion and peer review) and PBL pedagogy. The interview results confirm the most positive attitude towards the strategy of thinking maps. Finally, data from the interview also indicated that CT strategies enhanced their thinking skills, especially in Analysis and Evaluation.

Taken together, these findings contribute to the empirical study of combining CT teaching and PBL pedagogy in an EFL field and a CT-oriented course design to

improve students' thinking in the Chinese context. In particular, this study fills the gap between a unsatisfactory thinking learning outcome of the College English subject and CT requirement in the national educational policy. It is recommended for researchers to design a study with a control and experimental group to examine if positive results are achieved by integrating CT strategies into PBL pedagogy or the pedagogy alone. In addition, it is suggested for practitioners to create a learning community, increasing their CT awareness, acquiring related professional knowledge and developing thinking teaching literacy. Finally, it is advised for the policymakers or the leaders in HEIs, especially in the exam-oriented context, to reduce the pressure of the test results, increase the thinking teaching time and the percentage of measuring CT in the tests, eventually offering more opportunities for learners to practice and measure their thinking.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CBI	Content-Based Instruction
CE	College English
CET4	College English Test Band 4
CECR	College English Curriculum Requirements
CLT	Communicative Language Teaching
CT	Critical Thinking
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ESL	English as the Second Language
ESP	English for Specific Purposes
GE	General English
HEIs	Higher Education Institutions
L1	First language
L2	Second Language
PBL	Project-based Learning
PBLL	Project-based Language Learning
MOE	Chinese Ministry of Education
SCT	Sociocultural Theory
CSE	China's Standards of English Language Ability
ZPD	the Zone of Proximal Development

Chapter 1 Introduction

1.1 Introduction

Educators have emphasised the need for students to acquire critical thinking (CT) in tertiary education owing to its benefits for their academic study and future professions (D'Alessio *et al*., 2019; Huber, 2016; Li, 2016a; Snider, 2017). It not only allows undergraduates to understand intellectual knowledge deeply (Ennis, 2018) but also changes their ways of thinking and eventually makes a comprehensive career plan during the university period (Snider, 2017). Meanwhile, they are more likely to be creative and productive in the workplace with CT (Sasson *et al*., 2018; Soufi and See, 2019; Zou and Lee, 2021), such as making innovations in science or technology or handling international relationships to adapt to the trend of globalisation. Grigg and Lewis (2019, p.67) regard core thinking skills of "judgement and decision-making, deductive reasoning, complex problem-solving and system analysis" as five of the "top ten skills predicted for 2030 in rising occupations, such as hospitality, sports and leisure services" in the UK.

This research adopted a project-based learning (PBL) approach with embedded explicit thinking instruction in a College English (CE) course in order to examine its effectiveness in enhancing CT in a Chinese context. Specific to this chapter, it first introduces the rationale for and context of this study. Then, the significance of the research, followed by the research aim and questions, are presented.

1.2 Rationale of the Study

This study was inspired by my teaching experience of teaching students' writing in the CE course in China. They generally perform well in reading and listening, but not writing. This finding echoes some researchers' dissatisfaction with Chinese students writing performance (Dong, 2017; Huang, 1998). Besides low language proficiency, poor writing is evident in lacking logic structure. For instance, the examples learners provided may not support their arguments, or they just listed the examples without comparison or contrast, or they simply summarised their arguments without judgement. Some logical fallacies, such as causality or transition, also existed in their interpretations. Their thinking deficiencies in writing are confirmed by Hu (2017), who claims Chinese students usually concentrate on the writing structure instead of

using thinking skills, such as interpretation, analysis, evaluation etc. Even with a strong desire to improve writing in content and organisation, university students, especially in their first year, had no idea how to deal with it owing to the absence of relevant training in high school. Moreover, their insufficient thinking mismatches the new high-order thinking requirements of the curriculum reform in Chinese HEIs (The National Foreign Language Teaching Advisory Board, 2020). Given learners' poor thinking performance, the CE aims to help them develop higher-order thinking abilities. This was my initial motivation to conduct this research.

Many countries have recognised the CT significance by changing curriculums or introducing new policies. Soufi and See (2019) review education policies for undergraduates in Europe, the United States and Australia by clearly regulating it as a primary learning outcome as early as 2004. Some Asian countries, such as Singapore and Malaysia, have extensively promoted such education in recent decades (Pei *et al*., 2017). The Chinese government also publishes similar national educational guidelines to place a high value on CT cultivation as an essential tertiary teaching mission in China (Higher Education Law of the People's Republic of China 2018 Amendment, 2018). In addition, the latest round of the national curriculum reform in 2017 restresses the fundamental mission of HEIs is to cultivate undergraduates' innovation, CT and practical competency (Chen, B., 2017; Hu, 2017; Zhang *et al*., 2020). It was an official outlet for teaching CT in China. In order to be consistent with the requirement of the national educational policy and the curriculum reform, the revised College English Curriculum Requirements (CECR) adds CT requirements for non-English majors with three English levels, compared to the last version in 2016 (National Foreign Languages Teaching Advisory Board, 2020). This modification indicates that not only Chinese governmental administrators provide general policy support, but also policy-makers further offer a specific curriculum's CT guidance in the EFL field. Such policy and guidance supports are expected to increase the possibility of better CT outcomes theoretically.

However, a gap exists between a high expectation and an unsatisfactory outcome. To begin with, the revised CECR has clarified that the CE course should be an effective instrument to achieve a higher level of language proficiencies in listening, speaking, reading, writing and translating (The National Foreign Language Teaching Advisory,

2020). However, students' writing performance is far from satisfactory due to the underestimated writing instruction in EFL classes (Wang and Zhang, 2017; Zhan, 2013). Meanwhile, writing is one of the most challenging skills when learning a foreign language (Li and Zhang, 2021). This manifests in their comparatively poor performance in the writing section compared to the reading section in College English Test Band Four (CET4), a core national locally-developed high-stakes standardised test to assess non-English majors' English proficiency in China (Bai, 2020; Fan and Frost, 2022; Li, 2017, Qian and Cumming, 2017). A similar situation occurs in IELTS. The mean writing scores of students from mainland China in 2019 still lag behind the global average level, which is at a disadvantage compared to the reading and listening part (British Council, 2020). The comparatively low mean scores in writing suggest that Chinese students need to improve their writing. Performance in those tests may not be served as the only channel to assess writing abilities. However, the average grades in the above two large-scale tests can still be a valid index to draw a general portrait of students' comparatively weak writing performance. In addition, integrating CT into language classrooms has been peripheral, particularly in the EFL field (Li, 2020). Given this insufficient CT teaching practice, researchers or practitioners should strive to explore innovative instructional approaches and testify to the feasibility and effectiveness of improving learners' thinking in different EFL courses.

In addition to reflecting my personal teaching experience and national education policy support, the rationale also lies in the socialcultural theory (SCT) in general, which has a growing influence on understanding the relationship between English language learning and CT development (Chen, 2017; Li, 2016b, Mercer *et al*., 2019, Wegerif *et al*., 1999). As the core construct, Zone of Proximal Development (ZPD) is specifically utilised due to its benefit of creating a collaborative and interactive setting for students (Li, 2020), especially the ones who are unable to accomplish thinking tasks independently, to internalise and then externalise the necessary knowledge in order to fill the knowledge gap and eventually improve their thinking. This theory is claimed by many studies, believing an explicit way is the most effective to teach CT in language teaching classrooms (Abrami *et al*., 2008; Higgins, 2015; Lin, 2014; Lin, 2018; Tsui, 2002). It means CT instruction should be infused with the requirement of a curriculum with a clear teaching purpose and thinking activities or strategies. As

mentioned, thinking development has been officially listed as a new teaching objective with a brief description based on the revised CECR in 2020. However, merely listing a teaching objective without further elaborating on a course design and implementation and professional CT knowledge may still undermine students' thinking enhancement (Abrami *et al*., 2008). Thus, teachers must explore specific CT knowledge and pedagogy to help students foster thinking. Given the negative influence of the current dominating teacher-centred pedagogy in Chinese EFL classrooms (Gu and Schweisfurth, 2006; Tian, 2008; Tian and Low, 2011; Wang and Seepho, 2017; Zhang *et al*., 2020) and its dissatisfactory CT learning outcomes (Zhang *et al*., 2020; Zhao, 2012), a democratic, student-centred teaching method with explicit instructions is expected to help students improve thinking abilities.

Recent studies report that PBL, a student-centred pedagogy, outperforms in stimulating students to think critically than a teacher-centred traditional instructional method (Handhika *et al*., 2018; Kavlu, 2017). At the same time, it also improves students' language proficiency (Simpson, 2011). Beckett (2002) supports these ideas by stating its contribution to students' learning. Namely, more opportunities are provided for students to acquire content knowledge deeply and improve pertinent thinking skills, such as problem-solving, cooperation, independence and decision-making.

Integrating PBL into EFL courses under the guidance of the English curriculum aimed at developing CT is assumed to be applicable to the Chinese teaching context (Wang, 2020). On the one hand, lecturers are required to teach intellectual knowledge through standardised textbooks. Students could acquire and exercise content knowledge by completing different projects. On the other hand, teachers are encouraged to use innovative pedagogical instructions to cultivate students' high-order thinking skills to meet the demands of the curriculum reform (The National Foreign Language Teaching Advisory, 2020). Unlike traditional PBL teaching without embedding CT instruction elements, this study explored embedding a clear CT teaching objective and three CT strategies in the course design to provide students with more opportunities to practice thinking during the project completion process. At the same time, thinking development may allow students to understand the content knowledge better. Equally important, this method may fill the research gap in exploring practical CT

instructional approaches, owing to teachers' lack of and demand for CT professional training (Haas and Keeley, 1998; Zhang *et al*., 2020). In this light, PBL teaching intervention is anticipated to link subject content knowledge and thinking. Moreover, the possibility and effectiveness of situating CT teaching in a specific subject have been proved by several studies (Bağ and Gürsoy, 2021; Cáceres *et al*., 2020; Tsui, 2002).

1.3 Context of the Study

This study was conducted with first-year university students in an EFL course-College English (CE). On the basis of the revised CECR, its primary teaching objective is to enhance students' English application abilities and increase their intercultural communication awareness and competence. Meanwhile, it also aims to develop learner autonomy, raise comprehensive cultural literacy and cultivate the humanistic spirit and critical thinking capacities in order to meet the needs of academic study, social communication, and a future career (The National Foreign Language Teaching Advisory, 2020). Specific to CT learning outcomes, the CE course in this study seeks to allow students to analyse, interpret and evaluate information from different perspectives, understand the issues behind different cultures, and eventually achieve effective intercultural communication.

1.3.1 Issues in EFL teaching in Chinese HE

English is one of the most popular foreign languages studied in schools and universities in China. The important role of English has been growing since the Open Door Policy was introduced in 1978 as China continues to open up to the world and expands its influence in global markets. The importance of English in China primarily arises from its recognition as the language of international business, science, technology and higher education. As a result, English is studied as a compulsory subject by learners from primary schools to universities. In HE, all students are required to pass a national English proficiency test (College English Test-CET). Although passing CET is no longer a prerequisite for a bachelor's degree, that result is still regarded as a common standard of language mastery in China (Chen and Zhang, 2019). In addition, English language proficiency is also becoming increasingly important because more and more Chinese undergraduates intend to receive further

education abroad, esp. in English-speaking countries in the globalised context. However, English teaching results are far from universally satisfactory (Wright and Zheng, 2016), given the substantial investment in teaching time and resources. Besides social factors of the large class size (Wright and Zheng, 2016) and the actual limited use of English (Wei and Su, 2015), some scholars attribute this dilemma to the influence of culture, Confucianism (Deng, 2014; Fang, 2018, Liang and Fung, 2021), which advocates the authority of the teacher and encourages students to acquire knowledge by rote learning or memorization, rather than to construct and create knowledge (Li, 2016). It is not only detrimental to the improvement of language proficiency, but also the CT development, which is further detailed explained in section 1.3.4.

Furthermore, the culture of learning in China places a great emphasis on education and academic achievement. The cultural emphasis on education has contributed to the popularity of English learning in China because many students pragmatically regard it as a way to gain an advance in their future careers. Together, the role of English and culture of learning have led to a widespread interest in English language across China. As a result, students from all levels are motivated to learn English and seek to improve their language proficiency.

Secondly, compared with the CECR without any CT description in 2016, the revised version in 2020 explicitly specifies CT abilities as one of four teaching requirements in five language skills: listening, speaking, reading, writing and translating. In addition, it elaborates three levels of requirements for each skill based on students' different language proficiencies. These CT descriptions of five language skills for three English levels provide EFL teachers with a clear teaching direction. This new added CT teaching requirement indicates education policymakers advocate CT teaching in EFL (Sun, 2017). Inspired by the policy encouragement and curriculum guidance, practitioners realise its significance and increase their instructional awareness (Zhang *et al*., 2020).

However, the revised CECR could not be tailor-made to meet the needs of different grades of students as a curriculum requirement for their degree program for four or five years. In addition, few studies have a precise or targeted CT definition for different language skills, such as reading and writing, and further explore operating

rules in the Chinese context (Hu, 2017), especially for non-English majors. In contrast, China's Standards of English Language Ability (the CSE), which is the first official national English abilities standard, are more suitable for satisfying the demands of students in different grades by describing nine levels of English abilities (more detailed information will be illustrated in section 1.3.3). However, few studies define and apply CT skills based on the CSE in theory or practice. It may be because CT requirements are implicitly scattered in different scales and sub-scales of the CSE instead of being separately listed. Therefore, it is challenging for practitioners to detect those latent and concrete thinking skills hidden in the CSE if they lack professional knowledge. It is reported that the most significant hindering factors for teachers are the absence of CT theoretical knowledge, concrete pedagogy and incapability in thinking (Zhang *et al*., 2020). As a result, CT instruction has not become commonplace in Chinese universities (Zhang *et al*., 2020). It is also far from satisfactory globally, especially in the details of the teaching intervention and its implementation for different CT requirements (Abrami *et al*., 2015; Cáceres *et al*., 2020; Dennett, 2014; Huber and Kuncel, 2016; Niu *et al*., 2013). In this light, research on teaching CT may still be at the theoretical stage (Djiwandono, 2013), and more empirical studies such as this are needed. Therefore, researchers should clarify what and how to teach and evaluate the possibility and effectiveness of the instructional approach. Specific to this study, it started from defining CT based on the thinking requirement of the CECR and the CSE.

1.3.2 College English

The College English (CE) is one of the compulsory modules for all non-English major undergraduates. It plays an irreplaceable role in talent cultivation in China (The National Foreign Language Teaching Advisory, 2020). The CECR has cancelled a mandatory regulation of the CE's credits' proportion- 10 percent- of a student's total credits of undergraduate study since 2016. It provides more freedom for universities to arrange English teaching, and CE still has the longest teaching period of three or four semesters among all the compulsory public courses. In this research, the CE course counts 14 (out of 129, 11%) credits in a student's total credits of undergraduate study, including three modules: General English (GE), English for Specific Purposes (ESP) and CET6 training sessions (Table 1.1). In addition, *New Horizon College*

English, which is a standard textbook published by Foreign Language Teaching and Research Press, was used. It covers various topics around students' life in university, such as education, love affairs, consumption concepts, etc.

Table 1.1 Three types of the CE course

Time allotment	The GE (9 credits)		The ESP (2 credits)		CET 6 Training class (online, 3 credits)
The 1st- 3rd semester (4 hours/week for each class)	The Reading and Writing class	The Listening and Speaking class			
The 4th semester (4 hours/week for each class)			The Medical English Reading class	The Listening and Speaking class	
The 3rd or 4th semester (3 hours/ week, week 4-10)					The Online CET6 Training Class

1.3.3 An requirement of English level for the first-year university students

CSE, referring to China's standard of English, defines three categories of English language ability-basic, intermediate and adcanced- with nine levels altogether. the English proficiency standard lays down detailed requirements for listening, speaking, reading and writing. It also translating, where standards are rare wordwide (British Council, 2022). It aligns with English proficiency standards in other countries as a local standard to assess students' English performance in China. The Chinese National Education Examinations Authority (NEEA) and the British Council (BC) have been devoted to linking CSE standards to the results in two tests of Aptis and IELTS and jointly released a technical report in 2019. Taking the requirement of the CSE and scores of the IELTS for example, the first and second year Chinese university students should achieve level five, equivalent to the overall score of 5.5 in the IELTS (Table 1.2).

Table 1.2 Results of linking IELTS and level 5 of the CSE (Dunlea *et al*., p.10)

CSE level / IELTS score	CSE 5
Listening	6

Reading	5.5
Speaking	5.5
Writing	5
Overall score	5.5

In the third or the fourth semester, all non-English major university students are expected to pass CET4, which is the most influential high-stakes English test in Chinese tertiary education (Qian and Cumming, 2017), to achieve Level 5 of CSE.

1.3.4 Chinese students' CT deficiency in EFL classes

Chinese international students' negative behaviour or poor performance in academic study in English-speaking countries, such as reticence in group discussions, unwillingness to answer questions in class, and a dearth of CT skills in academic writing, attracts many scholars' attention (Atkinson, 1997; Cortazzi and Jin, 1996; Durkin, 2008; Li and Wegerif, 2014; McGuire, 2007; Paton, 2005; Shirkhani and Fahim, 2011; Tian and Low, 2011; Wang and Seepho, 2017; Wegerif *et al*., 2014; Wu, 2019; Zare, 2015). The silent behaviour and poor thinking performance may leave a general impression that Chinese learners lack CT in academic study. It is often taken for granted that students should engage in class discussion to express their opinions (Cheng, 2000; Gu and Schweisfurth, 2006; Turner, 2006) and write critically according to the academic requirement in western countries' universities (Candon and Kelly-Riley, 2004; Zhong and Cheng, 2021). In comparison, Chinese students who studied in their home country rarely have the opportunities to practice thinking in their learning context and are also absent from CT (Gao, 1999; Huang, 1998; Liang and Fung, 2021; Zhang, 2016; Zhang and Kim, 2018). These scholars' studies on the comparatively low criticality of Chinese EFL students may make Chinese researchers reflect on and explore the causes of such deficiency (Hu, 2017; Wang and Seepho, 2017; Wen and Zhou, 2006; Zhang *et al*., 2020; Zou and Lee, 2021). These studies identified three factors, namely, higher institutional and cultural factors and low language proficiency, resulting in Chinese students' thinking deficiency.

In terms of the higher institutional factor, more and more practitioners have been making efforts to teach CT in their subjects to respond to a national curriculum reform requirement and policy from the Chinese Ministry of Education (MOE) since 2017 (Hu, 2017; Zhang *et al*., 2020). Unlike earlier studies claiming the ignorance of

teaching CT in China (Huang, 1998; Li, 2011, Tian, 2008), recent studies report a growing tendency (Cheng and Wan, 2017; Li and Liu, 2021; Zhang *et al*., 2020). Specific to the College Engish, the latest released English curriculum requirement (i.e. CECR) provides guidance for EFL teachers with a generic CT teaching objective as part of the students' degree programs for four or five years. Additionally, the first published national standard for assessing English level in 2018 (i.e. CSE) offers knowledge support to university students with fine-grained CT skills requirements and different grades. Students are expected to exercise thinking skills at different learning stages, increase thinking awareness, and eventually become qualified critical thinkers when graduating from university. However, CT descriptions in the CSE are scattered in various sections rather than being listed in a separate section. It is challenging for teachers to select such implicit CT skills from the CSE and embed them in teaching practice, especially for those who lack of CT professional knowledge.

Even with higher motivation and enthusiasm for teaching CT, quite a high percentage of faculties in HEIs lack a deep understanding of its nature (AlKhoudary, 2015; Pithers and Soden, 2000; Snider, 2017, Zou and Lee, 2021). Teachers may face difficulties in teaching theory (Stapleton, 2010; Yuan and Stapleton, 2020) or specific instructional strategies (Li, 2016b; Li and Liu, 2021; Soufi and See, 2019; Zhang *et al*., 2020). Lin and Xiang (2020, p.96) highlight the significant obstacles are "a lack of synthesized understanding about what it is meant by CT and what an integrated CT programme should involve." Being aware of these challenges, practitioners strongly support embedding CT in the curriculum and growing desire for relevant training (Soufi and See, 2019; Stapleton, 2010). Abrami *et al*. (2008) and Zhang *et al*. (2020) also confirm the necessity of active and purposeful training for teachers to achieve better CT teaching outcomes. Besides, CT skills are seldom tested or highlighted in exams in Chinese HEIs (Tian, 2008). If not assessed, students may be less motivated to acquire relevant knowledge due to their pragmatical attitudes and the pressure of an exam-oriented environment (Chen and Zhang, 2019; Tian and Low, 2011). The impact of the exam results on students' learning is typical in the Chinese cultural context (Cheng and Wan, 2017; Lee, 2018; Shi, 2009; Tian and Low, 2011). Therefore, no clear CT instructional guidance, teachers' insufficient professional

knowledge and peripheral CT measurement in exams result in a gap between a high CT demand and unsatisfied CT teaching practice and outcomes. For example, how to effectively teach critical thinking in university is still unresolved (Huber and Kuncel, 2016). In this sense, Chinese CT instruction is still at the policy level to some degree. Even for those western countries whose central task in higher education is to develop CT, the result has been far behind satisfactory (Flores *et al*., 2012).

Secondly, some scholars attribute students' weakness in CT to the negative influence of Confucian culture and a consequent passive learning context (Atkinson,1997; Chen, 2017; Durkin, 2008a; Wang and Seepho, 2017; Tian and Low, 2011). It is manifested in the worship of authority resulting in a teacher-dominated pedagogy (Gu and Schweisfurth, 2006; Ku and Ho, 2010; McGuire, 2007; Paton, 2005; Wang and Seepho, 2017) and students' passive learning style by memorisation (Durkin, 2008b; McBride *et al*., 2002; Tian and Low, 2011). In contrast to the western discussion-based and student-centred pedagogy, the expository teacher-centred pedagogy dominates Chinese classrooms, which values teachers' lectures to impart knowledge (Teng and Zhang, 2017; Tian, 2008, Wang and Seepho, 2017). It means students need to be obedient to authority (i.e. teachers) in class. This may explain why Chinese students tend to doubt themselves when holding opinions different from their teachers (Zhong and Cheng, 2021). Additionally, although memorisation may allow learners to acquire solid intellectual knowledge quickly, it is supposed to inhibit them from thinking independently (Paton, 2005; Tian, 2008) owing to the fewer opportunities to reflect on what they learnt (Li, 2011). These negative consequences discourage students from thinking critically.

The influence likewise impacts Chinese students' learning experience when studying abroad. Gu and Schweisfurth (2006) elaborate on students' shock in the UK classrooms when being encouraged to learn independently and engage in group discussions. Zhong and Cheng (2021) investigate students' perspectives of treating teachers as the authority when starting their studies in the UK. Both studies indicate that Chinese students' cultural background and previous experience (i.e. secondary school) influence their learning outcomes in English-speaking universities. Furthermore, many studies show that even though Chinese international students intend to participate in class discussions, they have less experience in giving proper

feedback when they are exposed to the discursive teaching method in west countries (Cheng, 2000; Clark and Gieve, 2006; Jones, 2005; Rastall, 2006; Spencer-Oatey and Xiong, 2006; Tian, 2008; Tian and Low, 2011). Therefore, they need to adapt to the new CT-required ambience, which tend to be highly motivated and employ ampler opportunities to exercise and eventually achieve higher-order thinking skills to meet the academic requirement.

Students' response to the CT-lacking and CT-required learning environments can be either a hindrance or impetus to cultivating critical thinking. Not every student has the opportunity to study abroad. Therefore, the most practical way to improve CT in a Chinese setting is to transform the current teaching mode from a teacher-centred to a student-centred class, offering students more opportunities to understand, analyse, interpret and apply what they learnt as independent thinkers. Teacher-centred pedagogy may serve well in Chinese history, but it seems not to work effectively in modern times with the changing teaching purpose. Namely, to help learners and their countries adapt to the radical transformation in economy, politics, ethics, and morality (Flores *et al*., 2012). Specific to EFL students, they need to grasp the 21st-century skills to solve complex problems (i.e., problem-solving, critical thinking, teamwork, and communication skills) based on some fundamental language skills (i.e., reading or writing skills) (Kavlu, 2017).

Finally, low English proficiency is believed to be another major barrier for EFL students to perform better in thinking (Bağ and Gürsoy, 2021; Cheng, 2000; Fong, 2003; Jin and Cortazzi, 2006; Kirby *et al*., 1996; Koda, 2005; Luk, 2012; Manalo and Sheppard, 2016; Robertson, 2000; Zhang *et al*., 2020; Zhong and Cheng, 2021). Some researchers even claim low language proficiency more greatly prohibits students from logically expressing their opinions than cultural influence (Floyd, 2011; Huang, 1998). Manalo and Sheppard (2016) further point out that EFL students require more recourses to cope with a foreign language cognition instead of using resources to express CT in comparison with native English speakers. Comparing the report writing between the first and second year of Japanese undergraduates, McKinley (2013) finds that the latter students with instructions in English evaluation language write more evaluative statements both in English (L2) and Japanese (L1) than the former without language training of critical thinking. It is suggested that students improve their

evaluation skills in academic writing tasks through proper language instruction. A similar view towards a close relationship between language difficulties and higher order thinking performance is shared by Floyd (2011). The results of students' CT performance in English (L2) and Chinese (L1) indicate that their performance in L1 is better than in L2, and it is more challenging for students to show their thinking in an L2 setting. In this sense, the English level indeed influences EFL university students' thinking performance. However, some teachers may misunderstand that learners cannot cultivate CT unless their English is good enough because language proficiency is the foundation of critical thinking (Zhang *et al*., 2020). Therefore, teachers need to grasp adequate CT knowledge so that they can design proper activities for learners with different English language levels.

Generally speaking, three factors result in deficient CT in this study: higher institutions, the Chinese traditional culture and English language difficulties. Inconsistency between educational policy requirements and teaching outcomes indicates CT teaching remains at the policy level. In addition, an exam-oriented learning context means that students are less likely to acquire CT knowledge. Even being inspired by teachers' CT instruction, a teacher-centred pedagogy and students' passive learning style owing to the traditional culture may also hinder them from fostering thinking skills. Finally, students' comparatively low English level also impinges on presenting thinking in tasks. Although the governments' demand for CT instruction in Asian countries is increasing, students' CT performance has been far behind satisfactory (Zhang *et al*., 2020). More empirical studies integrating thinking and EFL subject teaching need to be taken into consideration (Abrami *et al*., 2015; Niu *et al*., 2013).

1.4 Significance of the Study

As discussed before, there are inconsistencies between unsatisfactory learning outcomes of the CE curriculum and CT requirements in the national educational policy. Therefore, the significance of the present study lies in exploring an effective CT pedagogy of EFL courses, which is assumed to provide university faculties with a practical option for teaching CT.

Many studies have investigated the effectiveness of teaching CT skills in EFL classes for English majors in China (Gao and Wen, 2017; Hou, 2012; Lin, 2012; Liu, 2018; Liu and Jin, 2012; Ren, 2013; Sun, 2017; Zhou, 2015). However, few empirical research explores the impact of a student-centred pedagogy on students' thinking development in generic and subject content-specific thinking skills for non-English majors in a Chinese setting (Ding, 2016; Tian and Low, 2011; Zhou, Jiang *et al*., 2016). This research is the first few studies that define the content-specific thinking skills in writing based on the new published CSE (NEEA, 2018) and the latest version of CECR (The National Foreign Language Teaching Advisory, 2020).

In addition, as a comparatively innovative teaching approach originated from Dewey's problem method in education (Beckett, 2002), many researchers employ project-based learning to teach CT in different subjects in tertiary education (Kokotsaki *et al*., 2016). Specific to language teaching, researchers report potential thinking benefits for language learners in L1 and ESL/EFL courses. However, their studies focus on PBL's effectiveness in generic thinking enhancement from a holistic view (Kavlu, 2017; Rochmahwati, 2015; Sasson *et al*., 2018; Senthamarai and Chandran, 2016). Few PBL studies explore the specific improvement in core thinking skills and learners' attitudes towards the effectiveness of PBL in fostering their thinking.

This research proposes to bridge a gap between the CT theoretical research and practical thinking instruction in the Asian EFL context (Snider, 2017). Studies have proved the effectiveness of applicability to improve students' thinking through language teaching in L1 (Marin and Hapern, 2011) and ESL/EFL (He, 2017; Lin, 2014; Sun, 2017). CT studies in the context of Asian EFL higher education have paid considerable attention to two strands: different standards of understanding CT concepts in the east and west (Chen, 2017; Kim, 2003; Li, 2016a; Tan, 2017) and various teaching activities' impact on students' thinking performance (Harits, 2011; Harizaj and Hajrulla, 2017; Snider, 2017; Wang and Seepho, 2017; Yang *et al.*, 2008). Few studies implement explicit CT strategies to systematic teaching approaches and investigate their effectiveness. Therefore, it is necessary to explore whether CT-oriented systematic pedagogies (i.e. project-based learning) could be employed in Asian EFL classes.

1.5 Research Aims and Questions

Given the previous discussion about limited empirical research exploring how effectively integrating CT into university EFL courses and the effectiveness of the CT-oriented PBL approach in enhancing concrete thinking skills and students' attitudes towards it, this research sought to explore the applicability and effectiveness of such instructional method in English teaching at the university level by investigating learners' attitudes towards this approach and its impacts on students' CT development. Therefore, this research attempts to answer the following questions:

Q1. What is the impact of the CT-oriented PBL approach on EFL students' CT development?

Q2. What are the EFL students' attitudes towards and perceptions of the CT-oriented PBL approach in their EFL class?

Q3. How does explicit CT instruction affect EFL students' CT enhancement?

1.6 Outline of the Thesis

Chapter 1 starts with descriping the rationale and context of this study by raising the issues in EFL teaching, briefly introducing the College English module, participants' English learning requirement and Chinese students' CT deficiency. Then, the research significance is explained followed by a presentation of this study's research aims and questions.

Chapters 2 and 3 review the literature about critical thinking and project-based learning, including their own definitions, benefits, the effectiveness of thinking improvement, teaching practice and their relationships. Chapter 4 introduces how to utilise sociocultural theory to design a PBL course for thinking development and the conceptual PBL framework. It is followed by introducing the feasibility of CT strategies as mediation to teach thinking.

Methodology chapters outline the research design of the pragmatism as the methodological basis and the exploratory study with a single-case strategy using a mixed method to collect and analyse data (Chapter 5), review data collection instruments, procedures and analysis methods. and ethical considerations (Chapter 6) and discuss the reliability and validity of the research design (Chapter 7).

Chapters 8, 9 and 10 report the findings in response to the research questions. Firstly, students developed their thinking skills and disposition (Chapter 8). Secondly, students held positive attitudes towards the CT strategies and PBL intervention (Chapter 9). Finally, CT strategies helped students mainly develop two core thinking skills (i.e. Analysis and Evaluation), while students met the challenges of their restricted capacities and low efficiency of group work (Chapter 10).

Based on the previous three chapters' findings, Chapter 11 discusses how CT-oriented PBL intervention creates the context for thinking enhancement and analyses hindering factors. It also discusses the impact of PBL on students' CT learning outcomes, attitudes, and perceptions.

Chapter 12 reviews the research aims and three main findings. It is followed by contributions, limitations and implications for practice. Finally, recommendations for CT and PBL teaching are put forwards for future research.

PART I: LITERATURE REVIEW

On the basis of the research purpose of exploring the applicability and effectiveness of PBL in enhancing students' CT, it is necessary to review the literature on both sides, CT (Chapter 2) and PBL (Chapter 3). They are followed by introducing a conceptual CT-orinted PBL teaching framework in Chapter 4.

Chapter 2 Critical Thinking

2.1 Introduction

This chapter accounts for the CT definitions, the interweaving between CT and the English language. It is followed by illustrating CT instructional approaches and assessment instruments.

2.2 CT Definitions

2.2.1 Definitions in education

The CT concept can be traced back to Socrates 2,500 years ago. This Greek philosopher encouraged his students to recognise and regulate "confused meanings", "inadequate evidence", and "self-contradictory beliefs" by inquiring questions (Paul *et al.*, 1997, p.8). As the first scholar who came up with the term "critical thinking", American pragmatic philosopher John Dewey (1933) re-defined "reflective thinking" as "critical thinking", a name that he originated and persists until now in education. He stated the foundation involves reasoning, scientific enquiry and problem-solving, which means critical thinking is a learning process. Since then, Western scholars have reached a consensus about its significance in higher education. The global interest in CT started in the 1970s (Ennis, 2015). Asian countries, such as Japan, the Republic of Korea and Indonesia, shed light on their priority in the significance of CT and lay the stress on the student-centred instructional approach until 2013 (Law and Miura, 2015). Although the CT significance has been reached a consensus, there are diverse understandings of the definitions involving what constitutes CT and whether it includes a range of skills or dispositional tendencies.

On a rational basis, some researchers believe that critical thinking contains a series of higher-order thinking skills, such as analysis, evaluation and inference (Bensley, 1998; Cheung *et al*., 2002; Dale, 1991; Elder and Paul, 1994; Halpern, 1998; Norris, 1985;

Pithers and Soden, 2000; Siegel, 1988). Davidson (1998) criticises this view for oversimplifying the CT concept as a rational judgment while neglecting the prominence of its practical application to people's daily life. Likewise, Lipman (2003) argues this skill-based concept may ignore a cognitive thinking process and overemphasise CT results as activities or beliefs. This defect in definition may result in a single standard for assessing students' CT performance. Students could get high scores on CT tests but may not be able to apply CT skills to a real-life situation. In this case, merely obtaining high scores is meaningless for university students to learn thinking. In contrast to the perspective of a set of logical skills, Ennis (1987, p.10) defines CT as "a reasonable reflective thinking focused on deciding what to believe or do". Emphasising its cognitive process, he articulates another essential element, namely, "dispositions" besides "abilities", and defines these two components as "the Alpha Conception of Critical Thinking" (Ennis, 2018, p.166). As one of the most popular concepts, many scholars accepted his definition with different interpretations (Bagheri, 2015; Facione, 1990; Halpern, 1999; Sofo, 2004). These scholars also believe CT can be applied and transferred to specific disciplines as a generic skill.

Moore (2013, p.506) defined a disciplinary field as "judgement, skepticism, a simple originality, sensitive readings, rationality, an activist engagement with knowledge and self-reflexivity" after investigating educators' opinions on three subject areas of history, philosophy and culture. Even though his research is restricted to three subjects, he proves different disciplines have different standards for defining CT. In contrast to Ennis (1987), he believes CT is a set of subject-specific skills that cannot be transferred to other subjects.

Although there is much debate, scholars eventually explored consistent elements among diverse definitions to benefit thinking education. In 1990, 46 American professionals in high education published Delphi Report (Facione, 1990, p.3), providing a widespread CT definition, including critical thinking skills and thinking dispositions in education:

> *We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as an explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is*

> *based. CT is essential as a tool of inquiry. ... The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, 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. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society.*

This statement highlights it as a dual-dimensional concept including a cognitive aspect with six core skills of Interpretation, Analysis, Evaluation, Inference, Explanation and Self-regulation and a dispositional aspect with seven mindset attributes of Truth-seeking, Open-mindedness, Analyticity, Systematicity, Confidence in Reasoning, Inquisitiveness and Judiciousness (Appendix A). Thinking skills application involves an internalising process of inputting knowledge through interpretation, analysis and evaluation and an externalising process of outputting it by inference, explanation and self-regulation. This means critical thinkers can explain and evaluate the information they obtained and form, present and self-regulate it rationally (Sun, 2017). Besides cognitive skills, ideal critical thinkers should also possess certain personality traits. Facione (2015b) claims that these two dimensions are indispensable. "As water strengthens a thirsty plant, the affective dispositions are necessary for the CT skills identified to take root and flourish in students" (ibid., p.20). He further highlights that "being both willing and able to think critically" is essential for critical thinkers (2015a, p.3). In other words, students are expected to be both skilled (able) to solve problems or make judgements using CT skills and disposed (willing) to do so (Huber and Kuncel, 2016). Facione (2015a, p.4) simplified that consensus into one sentence definition:

> *"Critical thinking is using this process of purposeful, reflective judgement to decide in a thoughtful, truth-seeking and fair-minded way what to believe or what to do."*

This concise definition amplifies the elementary purpose of thinking critically: make a judgment. Besides, what follows to do is the next core thinking purpose, which is the

main distinction between this definition and other definitions. Critical thinking is more than a theory but an action guide. In this sense, core CT elements involve *what it is* as well as *how to do it*.

2.2.2 Definitions in EFL

In the EFL field, Báez (2004) combines CT concepts developed by Bloom (1956) and Facione (1990) and defines CT as seven skills: knowledge, interpretation, analysis, evaluation, inference, explanation, and self-regulation. This CT definition conforms to foreign language learning characteristics. It requires learners to constantly store, retrieve, translate, analyse, evaluate new linguistic knowledge, find the inner relationship between the first and second language, monitor learners' thinking process, and become critical thinkers with high language proficiency.

In EFL classrooms in China, Li *et al*. (2016, p.279) surveyed 473 EFL teachers' CT understanding and summarised some common elements, such as analysis or inference, to make reasonable arguments.

> *The ability to analyze materials, summarise, see from different perspectives, discover rules and patterns in language learning, make reasonable arguments with evidence, and apply language in real-life contexts.*

Li and Liu (2021) map their CT taxonomy in the Chinese EFL context by reviewing nine salient international frameworks, such as Delphi, Watson-Glaser, Ennis and Cambridge, and put forward five essential CT skills of analysis, inference, evaluation, synthesis and self-reflection/self-correction. Different definitions may contain various components. However, they shared the same elements: analysis, inference and evaluation. Different EFL subjects may stress them differently even with the same core thinking skill. Specific to EFL writing, Cottrell (2017) defines it as a reasoning process of giving reasons, using relevant evidence, comparing and evaluating alternative arguments, weighing up conflicts evidence and forming conclusions based on the evidence. Li and Liu (2021) claim that CT definition in EFL is much more complicated than a general CT concept; therefore, it is essential to clarify it with concrete skills or indicators considering the demand for solving particular problems in English subjects rather than everything. Consequently, exploring the key thinking components in different EFL subjects within a specific context is necessary.

2.2.3 Definitions in the Chinese context

As one of the most influential education pioneers, Confucius has considerably impacted the oriented countries for thousands of years. Recording his teaching philosophy, a classic book, *Analects*, elaborates in the relationship between thinking and learning from an oriented perspective and draws great attention to many contemporary scholars (Chen, 2017; Kim, 2003; Tan 2017). Confucius teaching philosophy contains two levels of learning. One is knowledge accumulation through lectures, and the other is new discovery by reflecting on what they acquired (Kim, 2013). Reflection is indispensable to the learning process as an effective way of discovering further information. This close connection reflects in his famous saying," 学而不思则罔，思而不学则殆". It means "Learning without thinking is a vain effort. Thinking without learning is a dangerous effort" (Kawson [Trans], 2000, cited in Kim, 2003, p.81). His highlight of reflection in thinking seems to be a salient difference compared to the western educators, who are keen on defining CT from a logical perspective, such as sound reasoning or what constitute an argument. Kim (2003, p.82) concludes and divides Confucian relfective thinking into two types according to Analects:

> *(a) relfection on the materials of knowledge, in order to synthesizes and systematize the raw materials into a whole, and to integrate them into oneself as wisdom; (b) reflection on oneself, first, in order to ensure that sythesis, systematization and integration proceed in an open-minded, fair and autonomous way, and, second, in order to integrate knowledge with the self, this internalizing it until it becomes oneself.*

Kim's interpretation underlines twofold meaning: reflecting on the knowledge acquired by others and themselves. Firstly, by reflecting on teachers' lectures and observing their behaviour, students could use the acquired generic knowledge (i.e. theory) to come to grips with their specific problems (i.e. practice). This purpose of reflection is similar to one of the aims of teaching thinking in the West: making a reasonable determination (APA, 1990). Secondly, students should clarify their accumulated knowledge by reflecting on themselves. Otherwise, they have no idea what knowledge they should acqurie next in order to perfect their knowledge system. To keep the idea in mind, one should always be humble when facing the unknown, be

open to other leanrners, and be sensible to one's strengths and weaknesses, resulting in a better understanding of knowledge, abilities and bias (Zhao, 2013). Such characteristics of learners as humility, open-mindedness and sensibility are essential for successful critical thinkers in the western educational system.

Another Confucian classic essay named *The Doctrine of the Mean* in *The Book of Rites* defines the relationships among learning, thinking and behaviour by describing a more fine-grained thinking process: "博学之，审问之，慎思之，明辨之，笃行之", which means people should study extensively, enquire accurately, reflect carefully, discriminate clearly and practice earnestly. In other words, the nature of education is to help learners have a deep knowledge base, encourage them to cultivate a questioning spirit, form a thinking habit carefully, and develop a capacity to distinguish between right and wrong so as to practice and implement their beliefs faithfully (the result of judgement) (Sun, 2017). This interpretation displays a cognitive thinking process and indicates Confucius's highlight of the significance of gaining enough knowledge (wisdom), which is treated as one of the premises to reflect and think critically. In contrast, knowledge accumulation seems not to be stressed in the western concept.

On the one hand, Confucius advocates the significance of thinking critically in the learning process. On the other hand, he barely focuses on specific thinking skills because he believes it is not valuable, even dangerous, to use reasoning skills to make determinations without considering the wisdom from preceding generations and everyday life experiences (Kim, 2003). In other words, as an essential component of Chinese culture, learners should not neglect the wisdom handed down from generation to generation. In this sense, the CT concept exists in a Chinese context with divergent emphasis or the framework of analysis systems compared to the western culture, which is natural to different cultures. Meanwhile, both CT understandings favour the importance of human emotional factors: how to be an authentic critical thinker and some common cognitive skills, such as self-reflection. Therefore, it is biased for Atkinson (1997) and Chen (2017) to regard it as a western cultural outcome.

2.2.4 The CT definition in this study

As discussed in section 2.2.2, CT definitions in different subjects lay diverse stress on the thinking components. Therefore, it is necessary to explore particular definitions in different subjects. Specific to the CT in EFL writing, Cottrell (2017) defines it as a reasoning process of giving reasons, using relevant evidence, comparing and evaluating alternative arguments, weighing up conflicts evidence and forming conclusions based on the evidence. This definition demystifies the reasoning process in writing. However, it lacks a reflective process, the ultimate and most thoughtful stage yet typically neglected in writing (Farahian *et al*., 2021). Therefore, besides the above reasoning process, this study's CT definition in EFL writing adds a reflective process to revise the writing. In this light, the thinking process involving reasoning and reflection manifests the nature of CT: producing a logical conclusion by reflection and making determination by analysing and evaluating multiple evidence (Ma and Luo, 2021). The thinking skills involved in these two processes align with the Delphi report's six core thinking skills (i.e. interpretation, analysis, evaluation, inference, explanation and self-regulation) (Appendix A). However, each skill's description in that report is highly generalised for most fields in education, and some may not be consistent in writing. The definitions of core thinking skills in writing are amplified as follows:

Interpretation: detecting the meaning or importance of data and clarifying them by paraphrasing to reduce the vagueness. For instance, to outline the characters, themes or ideas of a text to support an argument; identify the parts of speech or guess the meanings of new words with the help of syntactic knowledge.

Analysis: recognising the reasoning process among descriptions, information or opinions to support conclusions. For instance, to discuss topics with clear viewpoints and sound reasoning; make a comparison or contrast among different perspectives in argumentative texts.

Evaluation: evaluating the reliability of the claims and logical strength of the arguments. For instance, to comment on articles or chapters; articulate clear and convincing viewpoints; assess the extent of prior knowledge about the topic before reading.

Inference: drawing reasonable conclusions or making strong presumptions by detecting and verifying the evidence given all the possibilities. For instance, to collect evidence from different sources to support an argument; analyse argumentative texts on common topics to infer authors' implicit viewpoints and attitudes.

Explanation: stating the results of the conclusion or justifying the reasoning process by presenting the cogent arguments logically. For instance, to state reasons for supporting the claims and conclusion.

Self-regulation: revising statements or descriptions based on reflecting on the skills they learnt. For instance, to adjust the structure and content of compositions through students' peer review, re-organise the structure to revise the content based on others' comments.

Although the above definitions and core thinking skills descriptions in writing are essential, some practitioners may not easily operationalise (Li and Liu, 2021). In addition, English proficiency is vital to impact thinking in language expression, as previously discussed. Therefore, more targeted and fine-grained thinking descriptions are needed based on participants' language levels. As mentioned in section 1.3.3, first-year students in university are expected to achieve level five in the CSE. Accordingly, this study detected 15 thinking requirements in writing scattered implicitly in the CSE 5 (NLC, 2018). They were subsumed into six types based on the above illustrations of core thinking skills (Table 2.1), providing clear indicators for teaching CT in writing. Otherwise, my teaching would be "shooting arrows at a target as a general concept", even though it provided a concrete definition and core thinking skills descriptions (Mulnix, 2012, p. 464).

Table 2.1 Description of content-specific critical thinking skills in writing

Core CT skills	CT descriptions
interpretation	1. to justify or refute by paraphrasing or to cite other people's opinions. 能通过转述或引用他人的观点来进行论证或反驳。
	2. provide relevant examples to explain the key points. 提供相应的例子解释主要观点
Analysis	3. to elaborate on the main thesis statement 阐述中心论点
	4. to provide the main claims 能提供要点
	5. to provide the sub-claims 能提供次要点
	6. to compare or contrast 能对不同观点进行比较和分析

Evaluation	7. to assess whether claims or evidence are workable, valuable and convincing 评估论点或证据是否行得通、有价值或可信的
	8. to express the implications: discuss possible effects or results based on the statements 基于论证讨论可能的影响或结果
	9. to determine what needs doing in a given situation 判断在特定的情况下需要做什么
Inference	10. to collect evidence from different sources in support of claims (personal experience, statistics or citing sayings) 能收集不同来源的证据以支持自己的论点
	11. to use cohesive devices to demonstrate and construct logical and coherent relationships 能使用衔接手段表示对比、因果、递进或转折表示句子之间的关系，使表达具有连贯性和逻辑性。
	12. to draw logical conclusions 做一个合乎逻辑的结论
Explanation	13. to provide relevant reasons to support the key viewpoints. 提供支持主要观点的原因
Self-regulation	14. to correct linguistic errata and adjust the structure and content of compositions through student peer review 能通过写作互评修改语言错误，调整文章框架和内容
	15. to re-organise the structure, revise the content, and correct linguistic errata based on others' comments 能根据他人评语调整文章结构，修改内容及语言表达中存在的问题

In summary, the term "critical thinking" could be defined from diverted perspectives, and no complete agreement has been reached over its definition. However, some essential consensus on its element has been achieved in the following three aspects: firstly, the thinking purpose is to make a well-reasoned judgement; secondly, besides some good thinking skills, critical thinkers should have certain emotional qualities, such as having a passion for exploring the unknown, keeping alert to personal bias or keeping an openness to controversy. Finally, the thinking process should include meta-cognition, which means critical thinkers could use specific thinking skills to monitor, regulate and modify their thinking process.

2.3 Sociocultural Theory and Thinking Skills

Before discussing the interweaving relationship between CT and English language, it is necessary to present how thinking skills be shaped in language learning. This study presents its mechanism from a SCT perspective, which claims that mental development occurs in the mediation when learners interact with the environment (Wertsch, 2007). As a symbolic mediation, language is treated as both a cultural tool to develop and share the acquired knowledge with the learning partners, and a

psychological tool, a "tool for thought" (Li, 2011, p.147), to structure the processes and content of the individual thought (Mercer *et al*., 2019). Given the close connection between these two functions (Vygotsky, 1978), language learning is, therefore, regarded as social at first, and then individual process; inter-mental at first and then intra-mental development, embedding thinking in the ways of using language. In this sense, language teaching enjoys an unique advantage of improving learners' collective and individual thinking. Therefore, language classrooms can be regarded as "social learning communities" (Li, 2011, p. 147). In this study, students are encouraged to study in groups to accomplish their individual writing projects. Teachers' lectures and students' collaboration and interaction in class and after class create a space for learning, thinking and sharing. This space created by learning communities also fills the gaps between their inadequate abilities to accomplish projects independently and the assistances of teachers and their counterparts with higher abilities.

In addition, SCT theory also emphasises "cultural characteristics to the current circumstances in the historical context" (Cole and Gajdamaschko, 2007, p.195). Specific to thinking skills, people from different cultural backgrounds with diverse learning experiences may differ in their views and perceptions (Manalo *et al*., 2015). As discussed previously in section 2.2.3, western countries and China value CT in education in different emphasis. The former emphasises thinking as the reasoning process in the context of demonstrative and individualistic teaching environment, which stresses thinking as a personal achievement. In contrast, the latter emphasises the reflective thinking due to the influence of Confucian culture, which values relationships and responsibility. Considering the difference, the CT definition in this study involves a reflective element (see section 2.2.4) and accordingly, utilises a CT strategy (i.e. peer review) to help students increase such awareness. In addition, Confucian education featured in rote learning and obedience to authority is believed to be an adverse factor to develop thinking (see Chapter 1, section 1.3.4). In this sense, Chinese traditional culture indeed has a strong influence on students' thinking.

Due to the social and cultural impact of thinking development, this study follows the SCT to use student-centred PBL pedagogy, embedding collaborative CT strategies, such as peer review and group discussion, in different teaching stages to create a

collaborative and comparatively relaxed environment, with an intention to contribute students' thinking development. More detailed information will be illustrated in Chapter 4, section 4.2.

2.4 Critical Thinking and English Language

From the sociocultural perspective, higher mental development occurs in the mediation when learners interact with the environment. Language as symbolic mediation, is treated as both a cultural tool to develop and share the acquired knowledge with the learning partners, and a psychological tool, a "tool for thought" (Li, 2011, p.147), to structure the processes and content of the individual thought (Mercer *et al*., 2019). Given the close connection between these two functions (Vygotsky, 1978), language learning is, therefore, regarded as social at first, and then individual process; inter-mental at first and then intra-mental development, embedding thinking in the ways of using language.

This section starts with a discussion of the relationship between CT and the English language. Then it moves to explore the benefits of CT development for EFL learners, followed by illuminating the effectiveness of teaching CT in EFL courses, especially in writing.

2.4.1 Relationships between CT and the English language

Empirical studies find that English proficiency and CT performance are interconnected (Jiang *et al*., 2016; Li, 2020; Shirkhani and Fahim, 2011, Zou and Lee, 2021). In the first place, CT development contributes to students' better language learning performance. It is reported that English skills development (i.e. reading and writing) is in proportion to thinking enhancement (Pithers and Soden, 2000; Wilson, 2016; Snider, 2017; Tsui, 1999). One potential contribution of thinking is improving students' English expression abilities. Utilising cognitive skills, such as analysis, evaluation or inference, students may reflect on the differences and similarities in the expressive ways between their native language (i.e. Chinese) and English. Once make it, they "could probably make a judgement on how to use English to reflect its linguistic characteristics as a foreign language" (He and Lin, 2017, p.148-149). Moreover, learners are more likely to achieve higher language proficiency by presenting their perspectives logically or reasonably if they think critically and

creatively (Kabilan, 2000; Rafi, 2011). The other possible contribution of thinking is allowing students to use more appropriate language learning strategies. Bagheri (2015) proves that students with higher CT capacities intend to use more effective language learning strategies, resulting in better learning outcomes than those with fewer strategies. In this sense, successful language learners are assumed to have a higher potential to employ thinking abilities.

In the second place, as a cognitive tool, language allows learners to structure and organise their thinking process (Storch, 2017) and fostering linguistic proficiency positively impacts CT beyond language acquisition. As discussed above, students could use language more proficiently once motivated to apply thinking skills. Likewise, students' CT positive learning outcomes could be evaluated by how these thinking skills are better presented in their English works (Manalo and Sheppard, 2016). Through analysing 436 English essays, Jiang *et al.* (2016, p.14) conclude that Chinese EFL learners' higher English proficiency benefits analytical thinking from three aspects: "the ability of making distinctions, the degree of cognitive complexity and the degree of thinking activeness". Their conclusion indicates that English learning fosters learners' thinking abilities. Besides students' better thinking performance in language works or tasks, a more satisfying performance in CT tests is the second contribution of language proficiency (Floyd, 2011; Lun *et al*., 2010). Considering the close connection between language level and thinking performance and the advantages of language teaching to improve thinking skills, CT should be prioritised in EFL classrooms (Zhang *et al*., 2020). In addition, their interweaving relationship offers a natural opportunity to integrate CT into language classrooms (Luk and Lin, 2015). Therefore, it is possible to teach thinking in an EFL course.

Specific to EFL writing, students' poor writing performance is closely associated with their CT absence (section 1.2). In other words, their writing performance is directly proportional to CT skills. From a cognitive psychology perspective, writing serves as a way to externalise writers' thoughts (Oatley and Djikic, 2008). Lin and Xiang (2020) further explain that EFL student writers' primary issue is how to display their thoughts in an acceptable manner following language conventions. In this sense, through transforming their thoughts into words, their thoughts could be reflected on. Meanwhile, articulating their thoughts enables students to gain a deeper insight into

complex phenomena, resulting in a more reasonable problem solution (Storch, 2017). From an applied linguistic perspective, writing involves a CT process (AlKhoudary, 2015). When preparing for writing, students need to gather and synthesise ideas from multiple sources and determine the most related information to their topics, making clear what constructs their essays. During writing, they should consider how to organise such content orderly to clearly express their perspectives and make a logical conclusion, persuading readers to accept their opinions. After writing, students could correct linguistic errors and detect logical defects to refute their ideas, increasing their writing quality. This writing process of searching for, selecting, developing and modifying their ideas from a messy situation to a reasoned result involves thinking. In this light, learners with thinking skills, such as analysis, inference or evaluation, are more likely to excel at writing logical and well-organised compositions (AlKhoudary, 2015). Therefore, teaching CT in language classes could help students to think critically through writing.

Besides the internal relationship between CT and EFL writing, it is feasible for educators to measure whether students develop CT or what specific thinking skills development students make through content analysis compared with other English language skills (i.e. reading or speaking). In this way, writing seems to employ a unique advantage if researchers intend to research students' CT development.

2.4.2 Benefits of developing CT for EFL learners

McGuire (2007, p.227) lists ten benefits for individual learners to support critical thinking pedagogy by reviewing several available published articles:

(1). Provides one with a means of self-defense against manipulation;
(2). Promotes one's individual autonomy;
(3). Protects one against self-deception;
(4). Helps one to resolve ethical dilemmas for oneself;
(5). Enables one to take greater control of one's life;
(6). Enhances one's self-confidence;
(7). Increases one's intellectual independence;
(8). Improves one's linguistic skills;
(9). Increases one's persuasive power;
(10). Promotes well-functioning democracies.

For students studying abroad, developing CT abilities not only reduces the risk of being marginalised in the academic study due to insufficient knowledge about the universities' academic norms (Tian and Low, 2011) but also increases the possibility of gaining academic success. On the one hand, they are required to achieve a higher level of language proficiency. On the other hand, they should also express their opinions in written and spoken English based on universities' CT-focused academic criteria. From the viewpoint of Anglo-American universities, it is crucial to acquaint students with the CT concept by introducing critical reading and argumentative writing courses (Alnofaie, 2013; Paton, 2005). As a result, it would contribute to these international students adapting to foreign countries' academic learning environments and encourage these young learners to obtain academic success in the future.

Chinese univeristy students studying in their home country will also benefit from it. Once acquiring such knowledge and becoming more aware of CT, graduates are expected to utilise thinking skills with dispositions to be productive members of the workforce and qualified citizens of this society (Woodward-Kron, 2002). As the future mainstay of China, they will think, reflect upon and take actions to reform the current society, integrate into the international community and eventually achieve more effective communication with other countries.

2.4.3 Effectiveness of teaching CT in EFL class

As an increasingly prevalent instructional approach for Chinese EFL teachers, integrating CT teaching in EFL courses is considered to be an effective and time-saving method (Flores *et al*., 2012; Lin and Xiang, 2020; Moore, 2011; Snider, 2017; Tsui, 1999; Zhang *et al*., 2020). Empirical evidence also proves the effectiveness of higher-order thinking skills cultivation in a foreign language class (Ding, 2016; Song, 2016; Soufi and See, 2019).

Firstly, the CE course provides students with a platform to develop criticality. This compulsory subject running for two years in the universities allows teachers to design and integrate thinking strategies from one to four semesters flexibly and systematically. Secondly, most Chinese universities do not offer a separate CT course, and most lecturers have not been professionally trained in thinking. Therefore, it is assumed to be more feasible and permanent for teachers to revise their familiar and

common-used approaches to teach CT knowledge instead of adopting new but potentially challenging methods (Pithers and Soden, 2000; Snider, 2017; Tsui, 2002). Utilising familiar instructional approaches may save teachers' energy and release pressure from the CT teaching challenge. In addition, learning thinking in an EFL context enables students to heighten their awareness of investigating the surrounding world from different angles (Snider, 2017) by making reasonable judgments, especially when facing diverse views from diverse resources online. For example, learners will gain experience in how to choose convincing evidence from various channels to build their cognitive system and form their understanding of the world in this knowledge-explosion age. Finally, English teaching has its unique advantage in CT cultivation. Some language teaching activities are suitable for the interactive nature of CT teaching activities. For example, students could become familiar with the Socratic questioning conversation because it often appears in the discussion of English language courses (Snider, 2017). Therefore, integrating CT in language classrooms contributes to learners' CT and language learning (Zou and Lee, 2021).

Although EFL teaching may improve students' high order thinking skills, it does not always work unless using some effective teaching strategies (Abrami *et al*., 2008). Manalo and Sheppard (2016) suggest that students increase their critical language awareness by using proper pedagogy to teach specific thinking elements rather than utilising a generic approach without CT-targeted language teaching. In this sense, it is worthwhile to select a feasible CT pedagogy applied to the EFL course.

2.5 Critical Thinking teaching approaches

2.5.1 Teaching approaches in general

Besides the CT definition, the other debate lies in whether it could be developed as generic or discipline-specific skills. This divergence leads to an option for different pedagogies: teaching CT in an independent subject study or within the context of different disciplines. Based on their different standpoints, there are five CT instructional approaches (Niu *et al*., 2013). The first is the "general" approach, which means to teach CT separately from subject-matter learning. CT skills and dispositions are the teaching objectives instead of the subjects' contents. The second method is the "infusion" approach, which refers to teaching CT as an explicit goal within subject

matter teaching. The third method is the "immersion" approach, which refers to teaching CT as an implicit goal within subject matter teaching. The similarity of the second and the third methods is integrating CT with the specific subjects, and subject contents and CT teaching are the teaching objectives. The discrepancy is the former method has an explicit teaching objective while the latter has an implicit CT teaching objective. Abrami *et al*. (2008) and Cheng and Wan (2017) point out that the infusion approach of embedding CT skills into the subject content and explicitly listing them as a teaching objective works better than the immersion method with implicit instruction. The fourth is the "mixed" approach, which means teaching CT as a separate goal parallel with the subject matter teaching. Precisely, this method combines the general approach with either the infusion or immersion methods. For example, Ennis (2018) mixed a one-year specific CT training course with another one-year instruction applying these thinking skills to six subjects. The results are positive for CT development. The final approach is the "holistic" approach, teaching CT in students' all-inclusive subjects for the degree programs. For instance, Beijing Foreign Studies University in China has utilised this method for English majors' degree programs since 2011, including reading, writing, listening and speaking courses.

Abrami *et al*. (2008) assessed the extent to the effectiveness of the first four approaches, and all of them helped students to develop their CT significantly. The mixed method achieves the most significant effect. While the immersion approach achieves the slightest enhancement. In addition, moderate effect sizes are found in the general and infusion approach. Although the mixed approach is the most effective, integrating critical thinking into teachers' subjects seems universal (Cáceres *et al*., 2020; Zhang *et al*., 2020). It means practitioners mainly use immersion or infusion approaches in their teaching practice. Their preference for embedding CT teaching into subjects instead of teaching it in a stand-alone class indicates that the latter method is less feasible for most teachers, owing to their lack of CT professional knowledge, one of the teachers' primary challenges (Zhang *et al*., 2020). Therefore, teachers may find it difficult to apply the general or mixed method that needs to teach students' thinking in a separate subject in the teaching practice. Compared to the first four subjects, limited research is currently available on assessing the effectiveness of

the fifth holistic approach. Therefore, comparing its effectiveness to the other four methods is difficult. Even if it is reported to be more effective, educators may find it barely feasible to apply it to their teaching because it needs teachers to have CT professional knowledge. At the same time, the university should provide policy support to design the CT-oriented curriculum for the degree program. Therefore, the infusion is expected to be an ideal option for teachers who lack CT professional knowledge. This study plans to use the infusion approach by integrating CT into EFL teaching with clear CT teaching objectives and instruction to achieve better CT learning outcomes.

In summary, whether CT should be taught in a separate subject or a subject-matter course is not contradictory. Every discipline holds its specific theoretical basis, teaching strategies and goals. Therefore, intellectual knowledge and abilities required by a subject cannot be transferred easily to another on the surface. However, some generic skills, such as identifying the credibility of information, utilising reasoning and exploring problems from multiple perspectives, will be required for almost every curriculum in higher education. It means cultivating thinking skills can be achieved by combing special training and a discipline-based study (Sun, 2017). It echoes a growing trend of CT education by integrating teaching strategies into specific subjects (Cáceres *et al*., 2020; Zhang *et al*., 2020). However, more empirical studies are called for to explore the effectiveness of concrete strategies and how to integrate CT teaching into different subjects.

2.5.2 Teaching approaches in EFL

According to Atkinson (1997), the term "critical thinking" in language teaching emerges in the context of first language (L1) teaching. Then it has attracted scholars' attention to ESL/EFL teaching since the 1990s because more and more international students are going to English-speaking countries to study, and they need to apply CT in new and different ways (Afshar *et al.*, 2017; Zare, 2015).

Most EFL practitioners do not explicitly state which CT teaching approach they use among the five aforementioned approaches. Instead, they typically tend to use language teaching approaches such as a communicative approach (Zare, 2015), content-based learning (Du and Zhang, 2022; Liaw, 2007; Tanaka and Gilliland,

2017), or generic instructional approaches, such as problem-based learning (Ding, 2016) and project-based learning (Kavlu, 2017) to teach thinking in EFL courses. In addition, some researchers merely apply concrete teaching strategies or activities such as mind-mapping (Kaeppel, 2021; Khodadady and Ghanizadeh, 2011; Rahman and Ambreen, 2018), group discussion (Dallimore *et al*., 2004; Darani and Hosseinpour, 2019; Nussbaum, 2002), debating (Yang and Gamble, 2013; Zhang *et al*., 2020) or analytical questioning (Wang and Seepho, 2017) to EFL teaching. Few researchers applied thinking tasks, like the Odd One Out or Six Thinking Hats, to language teaching (Lin, 2018). No matter which teaching approach or activity they use, most studies report more or less positive CT learning outcomes. However, some researchers put forwards their suspicion of the positive result. For example, Zare (2015) questions that some group activities, like an impromptu discussion of free and spontaneous communication in the communicative approach, may decline the reflective exercises. Given that most EFL practitioners prefer to situate CT instruction within their subjects (Cáceres *et al*., 2020; Zhang *et al*., 2020), infusion or immersion approaches without designing separate CT courses are the primary forms they adopted in language classrooms. Their preference is in accordance with their greatest CT teaching challenge: a paucity of CT knowledge. In addition, it is more feasible for these language teachers to use their familiar methods or strategies to teach CT elements. Although teachers have different options in teaching practice, scholars suggest that an explicit instructional approach is the most effective (Snider, 2017; Soufi and See, 2019; Sun, 2017).

2.5.3 Thinking teaching in the Chinese EFL Field

Huang (1998) is the first scholar who claims English majors lack critical thinking in China. The manifestation lies in the statement in argumentative writings or the questions they need to ask in lectures. In many cases, when students are asked to express their opinions, they have no idea what to respond to. He suggests that EFL researchers should conduct more research to solve this problem. He also suggests researchers adjust the curriculum and reform teaching methods.

In response to Huang's call for further research, a number of scholars have increased attention to discussing the performance of or the reasons for CT deficiency (Gao, 1999; Wen *et al.*, 2014; Wen and Zhou, 2006) and conducting a wide-range survey of

students' thinking (Wang and Wen, 2011; Wen *et al.*, 2018), especially for the undergraduate English majors. CT teaching practice mainly applies to different courses of language skills, including reading (Hou, 2012; Zhou, 2015), writing (Gao and Wen, 2017; Liu, 2018; Sun, 2011) and speech modules (Lin, 2012; Liu and Jin, 2012; Ren, 2013; Sun, 2017). CT instruction in separate language classes has recently been extended to a comprehensive curriculum, especially in the universities of foreign studies. Since 2011, the Beijing Foreign Studies University, one of the most prestigious universities in China, has used the holistic approach to teach CT for English majors' degree programs, including CT teaching practice in different modules (i.e. reading, writing, listening and speaking), publishing a set of CT-enhanced textbooks in 2014, developing a TERRIFIC CT teaching framework (i.e. target, evaluate, routinize, reflect, inquire, fulfill, integrate, and cooperate) (Sun, 2019). This is one of the few teaching practices for undergraduates using a holistic approach in China. Du and Zhang (2022) further discuss the CT learning environment from students' perspectives at the same university following a new approach of LaCTIT (language and critical thinking integrated teaching) based on content-based language instruction for their degree program. Students in that study held strong positive attitudes towards the CT learning environment after four years of study. Although their teaching practice alleviates the adverse conditions for undergraduates majoring in English (Zhou, 2018), the outcomes are still far behind satisfaction in comparison with other majors (Huang, 1998; Wen *et al*., 2010).

Unlike some researchers' holistic CT teaching method for English majors, most scholars use immersion or infusion ways to integrate CT teaching in their EFL subjects for non-English majors. He and Lin (2017) suggest building a learning community for students to cultivate CT abilities in the CE course and for teachers to improve their CT literacy by gaining support from English teaching policies, curriculum design, and thinking training. Fu and Wang (2021) provide a blended teaching model for effective CT teaching in the CE course. They suggest teachers should break stereotyped teaching concepts, improve CT skills, set a clear teaching objective, optimise multiple teaching resources, and revise the evaluation system. Compared with studies for non-English majors, CT teaching for English majors seems more systematic because educators could spend the four academic years designing the

CT-oriented curriculum and develop different core thinking skills in different language courses. In comparison, CT teaching studies for non-English majors have long focused on a single CE course, such as a reading and writing course or listening and speaking course. However, few research design a comprehensive and systematic curriculum, including all the reading, writing, listening and speaking modules. Although more and more practitioners have been attempting to teach thinking, there is no consensus on ideal CT pedagogies (Zhang *et al*., 2020; Zou and Lee, 2021). This may motivate them to explore effective teaching strategies from diverse perspectives and theoretical and empirical support.

2.6 Critical Thinking Assessment

Abrami *et al*. (2015, p.286) summarise five CT assessment instruments:

> *(1). Standardised tests: These are well-established measures of CT or particular thinking skills and dispositions: The Watson Glaser Critical Thinking Appraisal Test (WGCTA), Cornell Critical Thinking Test, California Critical Thinking Skills Test (CCTST), and California Critical Thinking Disposition Inventory (CCTDI).*
>
> *(2). Tests/evaluations developed by a teacher: This category includes, for example, the content analysis of students' responses to interview questions and open-ended and essay-type tasks teachers used to address CT skills development in their students.*
>
> *(3). Tests developed by researchers (i.e., one or more of a study's authors): These are nonstandardized measures developed by a researcher for use in a particular study.*
>
> *(4). Tests developed by researchers who also taught the courses.*
>
> *(5). Secondary-source measures: These instruments are usually adopted from other sources with or without modifications. Researchers may use previously developed (standardized or unstandardized) instruments or modify them to meet the requirements of their research setting.*

The standardised test is thought to be more valid than the self-designed assessment because it follows rigorous development procedures and has been repeatedly validated by being introduced to the public. It usually evaluates the generic cognitive skills towards thinking. However, it fails to assess disciplinary-related or content-specific thinking skills. In order to overcome such a limitation and triangulate the CT results, I used two forms of assessment tools. The first was the standardised test. One was the CCTST, which aimed to test students' generic thinking skills development. The other

was the CCTDI, which intended to test students' thinking disposition development. Both of them were chosen as the following reasons. As one of the most commonly-used CT assessment instruments in the world, their reliability and validity have been testified by diverse studies in many countries (see Chapter 6, section 6.3.1). Specific to the Chinese context, the Chinese version of these CT tests is a longstanding tool used by scholars since 2001 (Luo and Yang, 2001) and its application is well-established for students in high school and university in China (Chen *et al*., 2010; Cui *et al*., 2021; Geng *et al*., 2012; Jiang *et al*., 2014; Miao, 2018; Yan, 2021; Zhu and Shen, 2004). In addition, thinking skills in this study were defined on the basis of the CT theory and classification in the Delphi report, which was the same as the theoretical foundation for designing CCTST and CCTDI. It means that the thinking skills students learnt in current study are in line with the tested CT content in these two tests and both of them are applicable in the Chinese context.

Abrami *et al* (2015) suggested a fourth CT assessment tool for researchers who are also teachers. In alignment with the suggestion, my second assessment tool was a questionnaire that investigated student attitudes towards the effectiveness of CT strategies and PBL teaching to improve their CT. This was then supported with a thinking rubric, which was designed based on the thinking descriptions of EFL writing in the CSE 5 and CT classification principles in Delphi report to test students' content-specific thinking skills' development in writing. Both standardised and self-designed assessment instruments are triangulated each other and eventually make this study more valid.

2.7 Summary

This chapter elaborated on what constitutes CT in education, EFL, and the Chinese setting. More importantly, it defined the specific CT requirement for the first-year students in this study according to the new published CSE. Next, it provides the theoretical baseline for CT teaching in language classrooms. It is followed by the interweaving of CT and English language, CT teaching approaches and assessment methods.

Chapter 3 Project-based Learning

3.1 Introduction

Having discussed CT definition, theory, the links between CT and English teaching, different teaching approaches and assessment tools in the previous chapter, the current study is expected to place a CT-oriented PBL teaching approach within the curriculum to cultivate thinking skills. Regarding PBL teaching, this chapter begins by stating the background information, providing different definitions in higher education, especially in EFL areas, explaining PBL teaching's pros and cons and then clarifying the potential connections between PBL and CT development, introducing various PBL teaching programs of EFL courses from a general to a Chinese setting. It is anticipated that PBL is a feasible way to help students increase their awareness and skills in thinking and language knowledge in EFL classrooms.

3.2 Definitions

The emergence of project-based learning (PBL), in the name of Project Method (Kilpatrick, 1918) or project work (Beckett, 1999), can be traced back to the 1920s. As an outcome of a progressive education movement, PBL encouraged teachers to employ student-centred and experiential approaches to educate students by providing activities involving problem-solving in the real world in the United States (Condliffe *et al*., 2017; Holt, 1994). David Snedden, the first expert, conceived its concept and applied it to vocational classes in science instruction. Then William Heard Kilpatrick, a follower of the great educator Dewey, made this teaching mode prevalent in general education and resulted in a curricular reform movement - the project curriculum (Holt, 1994). Although it has been criticised by some educators who advocate the significance of traditional teaching methods to develop intellectual knowledge, PBL has become gradually universal in all levels of education, from primary education to vocational and higher education, because it is believed to help students learn deeply and foster multiple essential skills such as thinking or collaborative abilities in order to gain success in academic study, future career and social life in the 21st century (Condliffe *et al*., 2017; Sasson *et al*., 2018). However, it employs disadvantages, like time-consuming (Handhika *et al*., 2018; Holt, 1994) or an absence of improving individual competence (Condliffe *et al*., 2017; Holt, 1994).

There are various interpretations of PBL from different perspectives. As a pioneer in the PBL field, Kilpatrick defined the Project Method as "an activity undertaken by students that really interested them" (Ravitch, 2000, p. 179). As a constructivist instructional approach (Benson, 2005; Saad and Zainudin, 2022), it is defined as "a student-centred form of instruction based on three constructivist principles: learning is context-specific, learners are involved actively in the learning process and they achieve their goals through social interactions and the sharing of knowledge and understanding" (Cocco, 2006, cited in Kokotsaki *et al*., 2016, p. 267).

In education, it employs different interpretations according to diverse subjects' curriculum requirements. Blumenfeld *et al*. (2000, p.150) define PBL in science as "the presumption is that students need opportunities to construct knowledge by solving real problems through asking and refining questions, designing and conducting investigations, gathering, analyzing, and interpreting information and data, drawing conclusions, and reporting findings". In the field of ESL/EFL, Beckett (2002, p.54) defines it as "a long-term (several weeks) activity that involves a variety of individual or cooperative tasks such as developing a research plan and questions, and implementing the plan through empirical or document research that includes collecting, analyzing, and reporting data orally and/or in writing". Similarly, Kavlu (2017) describes it as a step-by-step learning process: students display their final project outcome by planning, doing collaborative teamwork, evaluating, and reflecting on the steps, critical aspects and different tasks to improve linguistic skills and cognitive competencies. Unlike both definitions of Beckett (2002) and Kavlu (2017) highlighting a teaching procedure, Beckett and Slater (2018b, p.1) redefine it in the name of project-based language learning (PBLL), stressing teaching outcomes in an e-learning context: "it is a comprehensive approach that involves individual or group activities, such as research reports, web site development, and digital stories, and focuses on the development of language, content, and skills in an integrated and meaningful way." This indicates a shift of research emphasis from studying teaching processes to assessing teaching results in different areas. Although there is no consensus, Handhika *et al.* (2018), Kokotsaki *et al.* (2016), and Markham *et al*. (2003) summarise common PBL features:

-a student-centred teaching approach working collaboratively;

-aiming to acquire content knowledge and multiple skills to solve authentic and complex problems;

-completing an individual or a group project.

3.3 Advantages and Disadvantages

3.3.1 Advantages in general

On the basis of three domains of 21st-century skills of cognitive, intrapersonal and interpersonal competence developed by Pellegrino and Hilton (2012), Condliffe *et al*. (2017, p.35) follow the same testimonies into the PBL field. They define the cognitive domain as "competencies related to thinking skills, content knowledge and creativity, such as reasoning, problem solving"; the intrapersonal domain as "affective competencies, like self-regulation or flexibility"; the last interpersonal domain as "competencies used to express, interpret, and react to information, like communication or collaboration" (Appendix B).

Firstly, it increases students' engagement and motivation in learning (Handhika *et al*., 2018; Martinez, 2022; Sasson *et al*., 2018). Students are free to inquire about questions put forward by themselves rather than passively answer the questions raised by teachers or pursue proper solutions instead of finding the only correct answer (Gratchev and Jeng, 2018). Therefore, learners may become highly motivated or activated to study in depth and enjoy this investigating process, which makes their study more meaningful. Saad and Zainudin (2022) claim that students are inspired by this "learning-by-doing" method because of being engaged in studying actively.

Secondly, it helps students apply the content knowledge to practice (Lehmann *et al.*, 2008). This application process allows students to actively dig into connections between a theory and practice, assess potential risks, construct their understanding, find solutions to the challenges and eventually complete projects. This inquiry process of fully grasping subject knowledge by practical application enables their learning to be deep and meaningful (Leggett and Harrington, 2021).

Thirdly, it develops students' multi-competence in academic study and future careers (Lai and Guo, 2018; Martinez, 2022). Most current studies focus on the interpersonal domain like cooperative or communicative skills (Handhika *et al*., 2018; Kavlu, 2017;

Leonard, 1917, cited in Holt, 1994; Saenabl *et al*., 2018; Sasson, 2018). Being a student-centred instructional approach focused on teamwork, PBL provides learners with more opportunities to do project work through cooperation, division of labour or peer evaluation. When doing project work together, they listen to and exchange ideas and express their opinions clearly to their peers.

Finally, PBL enables individuals to explore their interests and abilities and makes them responsible for their learning. This active learning style overcomes obstacles of discouraging deep thinking in rote learning (Clark and Gieve, 2006; Praba *et al*., 2017), emphasised in a teacher-dominated pedagogy in Chinese education (Tang and Biggs, 1996).

In a nutshell, PBL helps students internalise the knowledge and skills acquired through learning activities, apply their understanding by carrying out projects in flexible forms (i.e. written reports, videos, posters), evaluate, and present projects in public (Kavlu, 2017). This planning-implementing-presenting process may allow university students to input knowledge and skills in class and output them to solve authentic problems in real lives. Therefore, the advantages of project work teaching may go beyond knowledge acquisition. More importantly, it involves a process of shaping thinking and an awareness of cooperation, contributing to the learner's future academic study and career path. Even though instructors or students may confront difficulties during PBL teaching and learning, the strengths outweigh the weaknesses.

3.3.2 Advantages in EFL Classes

Studies show that PBL yields better English learning outcomes than traditional pedagogies (Pellegrino and Hilton, 2012; Thomas, 2000; Zhang, 2013). The improvement lies in content knowledge (i.e. language skills) (Kavlu, 2017; Simpson, 2011), awareness of cognitive skills (i.e. decision-making abilities and CT skills) (Brophy and Alleman, 1991) and communicative and cooperative capacities (Alan and Stoller, 2005; Bas, 2011; Beckett and Slater, 2005; Cooper, 1995; Kavlu, 2017; Kettanun, 2015; Li, 2010; Rochmahwati, 2015; Simpson, 2011; Tricia, 1993). Stoller (2006, p.48) ranks the eight most commonly cited benefits from 16 publications about ESL/EFL teaching, and most of them are related to the cognitive domain according to the testimonies in section 3.3.1 (Table 3.1)

Table 3.1 The advantages of PBL for ESL/EFL learning and the according domains of 21st-century skills

Domains	Advantages of PBL (Stoller , 2006, p.48)
Cognitive	- Authenticity of experience and language
	- Intensity of motivation, involvement, engagement, participation, enjoyment, creativity
	- Enhanced language skills; repeated opportunities for output, modified input, and negotiated meaning; purposeful opportunities for an integrated focus on form and other aspects of language;
	- Increased content knowledge
	- Improved abilities to make decisions, be analytical, think critically, and solve problems
Intrapersonal	- Improved confidence, sense of self, self-esteem, attitude toward learning, comfort using language, satisfaction with achievement
	- Increased autonomy, independence, self-initiation, and willingness to take responsibility for own learning
Interpersonal	- Improved abilities to function in a group (including social, cooperative, and collaborative skills)

Specific to ESL/EFL writing, Ng *et al*. (2021) summarise several PBL benefits for by elementary and secondary, and tertiary students by analysing 60 research papers on PBL from 2004 to 2018. These advantages are listed as follows and subsumed into three domains of 21st century skills (Table 3.2).

Table 3.2 The advantages of PBL for ESL/EFL writing and the according domains of 21st century skills

Domains	Advantages of PBL (Ng *et al*., 2021)	
	Elementary and secondary students	Tertiary students
Cognitive	- Gains in academic achievement and test scores	- The acquisition of writing skills and other language skills
	- Improvement in sentence structure or the writing length	
Intrapersonal		- Increased motivation in language learning and learning to write;
		- Increased awareness or attention to form in writing, especially in online writing projects such as bogs and wikis. While in terms of function, it related to experiential, interpersonal and textual

		meanings
Interpersonal		

Based on the reviews of Stoller (2006) and Ng *et al*. (2021), the benefits of PBL in general ESL/EFL learning or ESL/EFL writing in particular focus on the cognitive and intrapersonal domain. Firstly, it helps students foster language skills and content knowledge simultaneously (Beckett and Slater, 2005; Moritoshi, 2017) and thus increase a meaningful language output (Brumfit, 1984; Stoller, 2006) during the learning-understanding-application process. Students are likely to be given more opportunities to develop English abilities by effectively interacting with each other around a topic in an authentic communicative context. In particular, they are expected to use English to share information, actively listen to peers' opinions, participate in discussions and finally reach a consensus when completing projects (Beckett, 2018a; Simpson, 2011). These interactions benefit students from language output involving writing and speaking abilities.

Secondly, it enhances students' cognitive skills (Alan and Stoller, 2005; Eslami and Garver, 2013; Martinez, 2022). For instance, learners develop their higher-order thinking in a problem-solving process when explaining their own opinion, inquiring and answering complicated questions, and evaluating and challenging each other (Markham *et al*., 2003). When doing the projects, students analyse, evaluate and make decision to choose the proper information. Meanwhile, after completing the projects, they reflect and self-correct or regulate their projects. In this sense, higher-order thinking skills contribute to better project learning outcomes. Thirdly, it improves students' social skills (Moritoshi, 2017). For instance, students enhance communicative or collaborative skills by supporting each other when encountering language or technological challenges or feeling frightened to do presentations (Foss *et al*., 2008; Kavlu, 2017). As the key component of PBL, group work provides students with more opportunities to listen to and learn from peers' ideas, cooperate with each other and examine their projects from multiple perspectives. As a result, they build up team spirit and gain mutual support, which is essential for their academic study and future career.

Finally, PBL also benefits students' electronic literacy cultivation within the digital age. Beckett and Slater (2021) in their latest volume about project-based language learning (PBLL) introduce several empirical research on and frameworks for technology-mediated PBLL with diverse students' levels and contexts. They conclude that the technology instruments have become essential for project implementation within the current ESL/EFL teaching and learning contexts. Their perspectives are in consistent with the prevalent trend of applying e-applications (i.e. software or social media) to doing, completing or presenting project work (i.e. classroom newsletters, video projects) (Beckett, 2018a, 2018b; Lee, 2014; Meyer and Forester, 2015; Sidman-Taveau, 2005; Zhang, 2015). It is believed that technology mediated PBLL makes learning dynamic, allowing students to present their projects in diverse forms (i.e. video, website), communicate with the peers from other regions even countries without limitations of time and space. In addition, embedding technology with PBL teaching may reduce the time students spend on completing projects, which could be one possible option to overcome one of the intrinsic defect: time-consuming. Finally, e-applications may increase students' interest and inspire their potential to be more actively engage in learning.

3.3.3 Disadvantages

Although PBL has many strengths to support students' study, researchers also have to consider the following challenges:

-It is time-consuming. If students are accustomed to a teacher-centred class, it will take a prolonged period to get used to this innovative student-centred teaching mode. In addition, they are likely to spend more time completing their projects through a completed process of problem-solving, collecting information and group discussion. Teachers may have to invest extra time preparing for the teaching design to satisfy the learning outcome (Handhika *et al*., 2018; Holt, 1994). As a result, it would overload both students and teachers (Gratchev and Jeng, 2018).

-Ambiguous activities. With less empirical research evidence, instructors may feel confused about what teaching activities are effectively employed in the classroom. Some activities probably provide limited opportunities for learners to authentically communicate or cooperate in gathering helpful information and completing the

projects (Alan and Stoller, 2005). This challenge indicates that teachers have to grasp relevant knowledge about implementing each stage, including the expected teaching outcomes (Megayanti *et al*., 2020).

-Lack of individual competence development. As a collaborative teaching method, PBL emphasises the significance of teamwork, which is likely to be criticised for neglecting students' unique differences and personal development (Condliffe *et al*., 2017; Holt, 1994). Therefore, teachers should consider the question of how to divide students into groups and accomplish project work according to their different learning abilities and language proficiency.

3.4 Project-based Learning and Critical Thinking

As a student-centred active teaching method situated in a situation of solving real-world problems, a positive PBL learning outcome and high-order thinking skills development are intertwined (Handhika *et al*., 2018; Megayanti *et al*., 2020), especially when compared to traditional approaches (Leggett and Harrington, 2019; Sasson *et al*., 2018). Students could both foster CT skills (i.e. reasoning process and problem-solving) when completing projects (Condliffe *et al.*, 2017; Shepherd, 1998, cited in Thomas, 2000, p.13) and develop dispositions such as truth-seeking, judiciousness, grit and flexibility (Condliffe *et al.*, 2017). With the development of thinking skills and dispositions, PBL allows students "to apply ideas, figure out how phenomena occurs, and solve challenging, real-world problems" (Krajcik and Shin, 2022, p.73). The learning outcomes are consistent with the core purpose of CT: developing competence in making reasonable decisions and solving problems. It implies that PBL and CT development are interactive teachings and learning processes, and PBL is a potential effective pedagogy for teaching CT. However, not all PBL teaching interventions achieve significant thinking learning outcomes. Unless teachers design projects allowing students to think deeply and reflectively, model how to think critically, and give students feedback on group discussions (Handhika *et al*., 2018).

Although PBL teaching and CT development are interwoven with each other, and most studies discuss the PBL effectiveness in EFL from a cognitive domain (section 3.3.2), few researchers have touched upon its effectiveness in improving critical

thinking (Beckett, 2002; Kokotsaki *et al*., 2016; Pellegrino and Hilton, 2012). Firstly, few studies infuse CT-enhanced instruction into project teaching and investigate the impact on CT development. It is more effective if the PBL course design follows a CT framework and CT-enhanced learning activities with an explicit thinking teaching objective (Sarioudin *et al*., 2015). Secondly, most PBL empirical studies in EFL report improving the generic CT skill (Handhika *et al*., 2018; Holt, 1994) rather than concrete core thinking skills. Finally, most research instruments for assessing the effectiveness of PBL in fostering thinking skills are self-reports, such as interviews, observations or questionnaires (Alan and Stoller, 2005; Beckett and Slater, 2005; Gibbes and Carson, 2014; Kavlu, 2017; Rochmahwati, 2015; Sasson, 2018; Stoller, 2002). Their results may be weak, even unreliable if the measurement merely depends on the self-reports.

Nevertheless, they are feasible for researchers under certain circumstances. Since students or teachers may only provide subjective information about what they believe happened (Cáceres, 2020; Thomas, 2000), it would be better to achieve valid results by collecting quantitative data. In this sense, the quantitative tools are needed to triangulate the results to increase the credibility and reliability of the study. Given the research limitations of PBL teaching in CT development, more empirical research is needed to investigate an more targeted and effective PBL teaching design to foster thinking, to investigate the specific thinking skills development based on PBL intervention and to explore more reliable ways to assess its effectiveness.

3.5 Teaching Practice

3.5.1 Teaching practice in ESL/EFL field

Since 1918, Project Method has drawn attention in English studies by Wilbur Hatfield, an *English Journal* editor (Beckett, 2002; Holt, 1994). In the middle of the 1970s, project work became a prominent English teaching pedagogy (Gibbes and Cartson, 2014; Stoller, 2006). In the 1980s and 1990s, PBL was introduced into ESL/EFL as an antithetical approach to the widespread teacher-centred pedagogy (Beckett, 2019; Beckett and Slater, 2018a). In addition, it usually acts as a natural enhancement or supplement to mainstream pedagogies (Stoller, 2006). Brumfit (1984) introduces project-based communicative language teaching in L2 classes under the

Communicative Language Teaching (CLT) theory. This integration of PBL and CLT enables ESL students to communicate in an authentic context (Beckett, 2002). Moreover, it creates a language learning environment to think critically to share, explain, reason, evaluate and challenge each other when preparing, doing and presenting projects (Jacobs and Farrell, 2003; Sofo, 2004; Zare, 2015). Alan and Stoller (2005, p.12-13) conduct PBL teaching based the theory of Content-Based Instruction (CBI) to carry out project work with a ten-step PBL framework in an ESL class:

Step 1: Students and instructor agree on a theme for the project;
Step 2: Students and instructor determine the final outcome of the project;
Step 3: Students and instructor structure the project;
Step 4: Instructor prepares students for the demands of information gathering;
Step 5: Students gather information;
Step 6: Instructor prepares students to compile and analyze data;
Step 7: Students compile and analyze information;
Step 8: Instructor prepares students for the language demands of the final activity;
Step 9: Students present the final product;
Step 10: Students evaluate the project.

Stoller (2006) positively comments that university students could choose various projects according to intellectual knowledge, motivating them to engage in learning compared to traditional classes. In addition, she claims teachers' scaffolds in steps 4, 6 and 8 support students in completing successful projects. Applying PBL in the content-based course provides a more flexible and engaging environment for EFL students in university to study collectively. Moritoshi (2017) generalises the above ten steps into four systematic teaching stages: the stage of planning (steps1-3), implementation (steps 4-8), show-and-tell (step 9) and evaluation (step 10). Similarly, Eslami and Garver (2013, cited in Mathews-Aydinli, 2007, p.3) list five procedures for teachers to design PBL teaching in EFL classes: "preteaching, introducing the problem and the language for solving the problem, grouping students and providing them with the necessary resources, observing and supporting the students, and following up and assessing their progress." These different PBL course designs on the

basis of language content are in line with students' learning rules and help teachers to achieve the language objectives

Based on constructivism, Beckett and Slater (2005) developed a project framework to help ESL/EFL students increase academic literacy, including linguistic skills, thinking skills, and intellectual knowledge. It contains two primary elements: "the planning graphic and the project diary" (ibid., p.110), allowing students to understand how valuable PBL is by writing down all the learning procedures, stimulating them to maximise project work's potential benefits and completing projects successfully. Slater and Beckett (2019, p.1) revise their framework by blending it with Mohan's knowledge framework about academic language to "create unit plans that explicitly integrate language, content, skills, and technology" for American graduate students. Students are given more opportunities and freedom to choose project topics they need for future academic study. This revised model is proven to be applied in ESL/EFL courses in an academic setting. All their studies indicate it is possible and practical to combine subject content and PBL in diverse EFL courses for different levels of students.

Although PBL teaching contributes to positive EFL teaching outcomes, the relevant studies on science and social subjects heavily outnumber language teaching and math. More empirical studies are called for to investigate whether and how promising project research is on EFL subjects (Beckett, 2002; Condliffe *et al*., 2017), especially cognitive skills development, for instance, decision-making, critical and problem-solving skills (Moritoshi, 2017; Stoller, 2006).

3.5.2 Teaching practice in the Chinese EFL field

Gu and Zhu (2002) are the pioneers in applying PBL teaching to Chinese EFL teaching in HEIs by conducting a teaching reform in 1997. They used multimedia-centred PBL instruction in a series of language-related subjects (i.e. Comprehensive English, English Reading and Business English) for English-majored undergraduates of Suzhou University. Gu (2007) expanded research participants from English to non-English majors to further investigate its effectiveness in 1999. It is a computer-assisted collaborative English writing project between Chinese and American undergraduate students under an authentic communicative learning context,

investigating PBL's impact on learners' learning motivation and writing performance and the roles of technology. Her research group claim many benefits from project work in an EFL setting, such as increasing knowledge curiosity, enhancing writing performance, developing general communicative competencies and raising awareness of cultural differences (Gu and Zhu, 2002). However, few Chinese researchers followed this novel teaching approach due to restrictions on the electronic teaching environment in most universities at that time (National Bureau of Statistics of China, 2002). Although meeting challenges in infusing technology with PBL, Chinese EFL researchers have still attempted to apply this western-originated teaching approach in different courses, such as CE (Gao, 2010) and ESP (Wang, H., 2013). Wang (2020) reports an increasing trend for PBL instructions in language classes from 2002 to 2017, indicating a rising research interest. As the first scholar to explicitly localise PBL in China, Zhang (2015) developed a framework of innovation-oriented project-based learning (iPBL) based on his six years of PBL teaching practice in the course of English Through Projects for third-year English majors at Nankai University. He claims that students enhance innovative capacity and critical thinking without presenting valid evidence.

Wang (2020) reviews 39 PBLL studies published between 2002 and 2017 and concludes three common features in the Chinese EFL field. Firstly, the research quantity has increased, especially in the past decade. Whereas the quality needs to be strengthened by more valid assessment instruments. For example, most studies employ qualitative methods, such as interviews, class observation, or reflective diaries. These qualitative data may provide detailed, specific information or different perspectives to help scholars gain deep insight into this teaching process. However, quantitative data is called for to describe a general trend or provide a supplementary numerical interpretation of the teaching outcome to triangulate the qualitative results.

Secondly, when discussing the effectiveness of PBLL, current studies focus on the macroscopic level, such as the effects on motivation (Gao, 2010; Wang, 2013), collaborations (Gao, 2010; Shi, 2009) or autonomy (Deng and Wang, 2009), and students' experiences, such as their perceptions (Yu, 2017) or satisfactions of the learning outcomes (Wang, 2012), and their efforts in PBL (Xia and Zhang, 2017). Few studies focus on a microscopic level. For example, enhancing particular skills,

such as language acquisition (Chang *et al*., 2019; Gu, 2007; Gu and Zhu, 2002; Yang and Han, 2012) and no study has been found on fostering critical thinking. In addition to few PBL teaching for non-English majors in the CE course (Gao, 2010) or EAP/ESP courses (Gu and Ye, 2017; Li and Du, 2014), more applications are for English majors, such as English Writing (Gu, 2007; Gu and Zhu, 2002), English Newspaper Reading (Yu, 2012), Studies of English-speaking Countries (Chang *et al*., 2019) and Chinese Culture Teaching (Gu and Ye, 2017). That is to say, PBL teaching in EFL courses in HEIs is mainly used for students with higher English proficiencies. The above features and trends indicate that PBL instructions are still at an initial stage. More empirical research with higher qualities is needed to target non-English majors to assess its effectiveness.

3.6 Summary

This chapter first described the emergence, PBL definitions in general and specific to EFL teaching and its cons and pros. The connection between PBL and CT was then presented to explore the possibility of cultivating CT in PBL classrooms. Finally, it introduced PBL teaching practice in EFL subjects in general and in Chinese tertiary context in particular.

Chapter 4 Framework of the Study

4.1 Introduction

The review of critical thinking in chapter two and project-based learning in chapter three has shown the strong connections between CT and the English language and between CT and PBL teaching, and PBL is believed to be an effective pedagogical approach to improving students' thinking skills in EFL teaching practice. However, it does not mean that any PBL pedagogy automatically enhances thinking. Students may not achieve expected results if they are simply divided into several groups to design or create something unless projects have been designed with an explicit CT teaching purpose (Mergendoller, 2012) and substantial deliberate practice (Mulnix, 2012). Therefore, it is necessary to initially explore the theoretical motivation for designing PBL pedagogy and discuss how this course design achieves the CT learning outcome. It is followed by a brief introduction to three CT strategies (i.e. thinking maps, peer review and group discussion) by discussing the possibility of embedding them with different stages in this pedagogy.

4.2 Sociocultural Theory and CT-oriented PBL Course Design

CT teaching in this research was designed through the lens of sociocultural theory (SCT), is believed to be the source of deep learning and development in social and cultural interaction instead of in an individual's mind (Panhwar *et al*., 2016; Swain *et al*., 2015). This interaction is mediated by a third factor (mediator) of material tools (e.g. textbooks, computers) and/or symbolic signs (e.g. language, cultural concepts) (van Compernolle and Williams, 2013). Both material and symbolic tools serve as mediation, a central principle of SCT. Vygotsky (1978) and Wertsch (2007) claim that learners' higher mental development depends on the presence of mediation in their interaction with the environment. Symbolic mediation contributes to not only "quantifiable progress in terms of competence and fluency" but also "cognitive improvement in terms of critical reasoning and thinking" (Lantolf and Poehner, 2008, cited in Panhwar *et al*., 2016, p.184). This mental process involves two phases. The first is the mediation stage of acquiring a symbolic artefact in a particular cultural environment as an external symbolic tool (i.e. language) (Kozulin, 2018). This stage occurs when learners acquire new knowledge by interacting with the contextualised world outside of their minds. The second is the internalisation stage of transforming

the external symbolic tool into a psychological tool (ibid., 2018). In this stage, learners "voluntarily organise and control (i.e. mediate) mental activity and bring it to the fore in carrying out practical activity in the material world" (Lantolf and Thorne, 2006, p.62). It implies that internalisation enables students to master a problem-solving ability, thus achieving independent learning as part of the inner mental process. Therefore, if practitioners intend to develop students' thinking, the pedagogy should place students in the central position and involve their learning in two stages: acquisition and internalisation.

As a student-centred pedagogy, PBL involves the same two mediation stages, encouraging learners to solve authentic problems using acquired content knowledge and multiple skills to complete projects collaboratively. It aligns with SCT, which treats students as active learners who recreate, modify and extend the knowledge in and through collaborative knowledge building and individual understanding, rather than passive knowledge receivers (Storch, 2017). In this sense, it is supposed to be more effective than the traditional teacher-fronted pedagogy for cognitive development. However, not all PBL teaching intervention results in a significant thinking transformation since different forms of mediation differentiate the level of cognitive development (Kozulin, 2018). van Compernolle and Williams (2013) suggest mediation can be intentionally embedded in teaching for pedagogical purposes. It implies mediation could be introduced for developing CT in this study. Therefore, CT instruction as mediation should be added to PBL teaching to achieve significant cognitive enhancement.

What is the optimal CT instruction contributing to students' learning? A well-known concept of Zone of Proximal Development (ZPD), developed by Vygostky (1978), answers this question by explaining that "the zone focused on the relationship between instruction and development" (Chaiklin, 2003, p.39). Specifically, development occurs in the interaction between teachers and students, especially the guidance or help for those who cannot complete the task by themselves. In this sense, ZPD fills the gap between learners' lack of abilities to carry out the tasks independently and the assistance of more experienced others (Panhwar *et al*., 2016). As human mediation, teachers play an indispensable role in this gap-filling process. They should clearly explain teaching contents to students, especially when imparting

new knowledge (Lin, 2018) and explicitly show good CT examples to students (Mergendoller, 2013; Moon, 2008). In this study, due to students' absence of CT training and CT competence in their previous learning experience, teachers' lectures about CT knowledge based on the content knowledge of the CE are needed to fill the gap initially. Through lectures, students are expected to acquire the content language knowledge and thinking skills, which are the premises for completing the given task (i.e. writing projects).

Besides CT guidance provided by lectures, the interaction of collaboration between teachers and students provides "support at the moments when maturing functions are inadequate" (Chaiklin, 2003, p.54). Unlike the ideal situation of ZPD that scaffolds novices' (i.e. students) learning through experts (i.e. teachers), this study expands the guidance between teachers and students to the support among students within a group out of three considerations. Firstly, the comparatively large class size (10 groups of 50 students per class) may not allow teachers to interact with every student and provide instant support in class. Secondly, students in this study are top of the university with the highest mean scores in Gaokao (it is discussed further in Chapter 5, section 5.3.3.3). It means they have comparatively high learning abilities. Some may understand better when acquiring new knowledge and become potential "student experts" who can support other inexperienced peers. Finally, accustomed to the teacher-authority traditional culture, asking for teachers' help is the second choice for some students due to their fear of or reluctance to communicate with teachers. Instead, seeking help from peers is less stressful and comfortable. In this sense, the interaction in this study occurs in the contextualised assistance between expert members of the community (e.g. teachers, knowledgeable students) and novices (e.g. less knowledgeable students).

Appropriate assistance could promote cognitive development but also measure the extent to the development. Students are judged to have greater potential for cognitive development if they can carry out the task with assistance than the ones who cannot take advantage of it (Storch, 2017; Vygotsky, 1978). In addition, teacher-student and student-student interaction is the central section in language classrooms. It means teachers and students are encouraged to use strategies collaboratively to engage in the

language learning environment (Vygotsky, 1978). Therefore, students need to work in groups to be supportive.

In this research, every student has to organise and complete a well-structured writing project with at least 500 words. Because they have never done it before, this task could be challenging, and they need assistance from the teacher and peers. Storch (2017) suggests that the offered support not only needs to make challenging tasks accessible but encourages students to engage in learning. Therefore, all the support acts as the scaffold to assist learners in accomplishing the task beyond their present capacity, step by step and finally complete it independently once employing enough competence with assistance. In this sense, the strategies should be conducted collaboratively to help students complete writing projects. Collaborative thinking maps, group discussion and peer review activities are selected to help students apply their acquired language and CT knowledge to complete writing projects in the corresponding pre-writing, drafting and revising stages. Meanwhile, these activities are effective language learning and CT training strategies for students to improve their thinking performance in writing (more detailed information will be illustrated in the following sections 4.4, 4.5, 4.6). In brief, a core SCT concept of mediation explains why PBL teaching outperforms the traditional teaching method and why CT instruction should be embedded in it. The other well-known concept of ZPD responds to the question of what CT instruction should use.

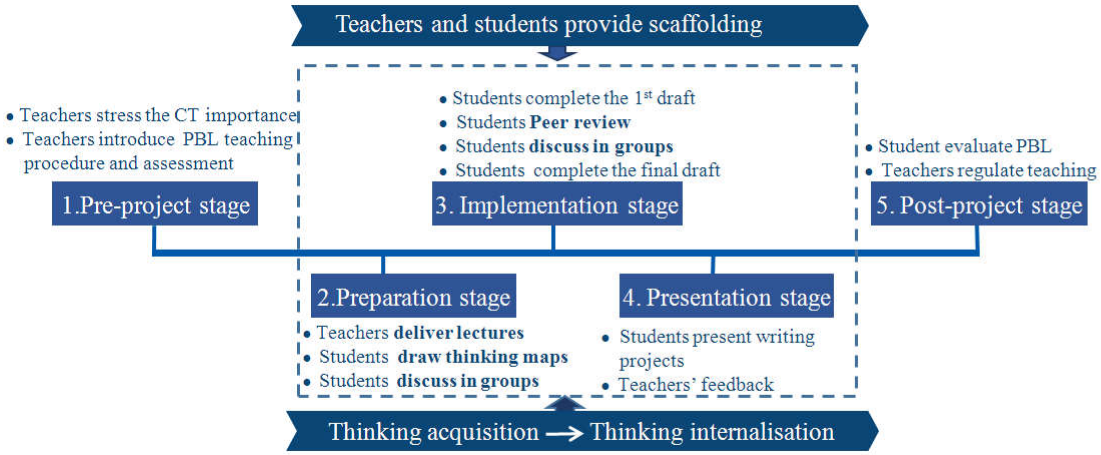
4.3 Conceptual Framework of Current Study

This study developed a new framework by adopting Alan and Stoller's (2005) ten steps of project work in EFL teaching (it will be illustrated in Chapter 3, section 3.5.1). Unlike their general framework of developing diverse projects for maximising PBL benefits in multiple dimensions (i.e. content knowledge and language mastery), this study focuses on writing projects and CT enhancement in particular. Instead of merely providing language support, the major alteration of this framework is adding steps of CT instruction to serve the above two key points, including lectures on content-specific thinking skills and language demands and three strategies that can improve both linguistic proficiencies and CT in order to acquire and internalise thinking knowledge when learning English (Figure 4.1). The other alteration is removing three steps in Alan and Stoller's framework of information gathering,

compiling and analysis because data processing is not much involved in non-academic writing. This new framework focuses on acquiring and internalising CT in the learning process on the basis of SCT.

Before presenting the conceptual framework, briefly introducing the basic information about students' grouping and writing projects is necessary (more detailed information will be presented in Chapter 6, section 6.2). Participants (n=98) were divided into ten groups with five members for each group to study collaboratively. Everyone should complete one writing project independently and revise four essays collectively. Meanwhile, they can choose any title within the one out of five topics they are interested in according to the five units of the textbook.

Figure 4.1 A framework of CT-oriented PBL course design



Due to students' lack of PBL and CT learning experience, this framework begins with orientation (pre-project stage) to clarify the CT teaching objective and introduce the critical information of this module. The introduction (i.e. CT significance, PBL benefits and teaching procedures) allows students to clearly understand what they should learn and detect potential gaps between learning requirement and their language and thinking abilities, leading to a directed effort. This stage is necessary for PBL start learners (Zhang, 2015) as a stimulant to help them engage in class and eventually lead to successful project work (Beckett and Slater, 2005, 2018a).

Next, teachers and students help the one who took charge of writing (project writers) prepare for writing at the preparation stage. Teachers deliver lectures, including linguistic knowledge (i.e. topic-related vocabulary, grammar and sentence patterns),

thinking skills in writing and CT strategies, such as thinking maps. All the presentations contribute to filling students' knowledge gaps and meeting the demands for completing writing projects. In this light, the lectures are regarded as part of the scaffolding (ZPD) to help students who lack language proficiencies and thinking abilities complete writing projects independently. To some degree, teachers' assistance and guidance of students are the most crucial element in carrying out and completing a successful project (Markham *et al*., 2003). From students' perspectives, they "appreciate that projects are mixed with traditional courses: they can apply knowledge studied traditionally, via the projects" (Rouvrais *et al*., 2006, p. 90). With the teachers' guidance, project writers determine a title and draw a thinking map before writing. Other peers discuss and evaluate the reasonability of logic and structure in that map through group discussion. Based on the discussion result, project writers judge what should be revised on the map. This mental process implies that mediation is done through symbolic tools of lectures (i.e. modelling how to analyse the passage's structure and draw a thinking map) and two strategies of thinking maps and group discussion. Learners acquired the demanded knowledge and internalised it by presenting the thinking process when creating their thinking maps and discussing the logical defects in groups. In this sense, as SCT indicates, mediation occurs in two phases of knowledge acquisition and internalisation when students, who are inexperienced in drawing the thinking map, can create it independently with the assistance of teachers and group members.

In the implementation stage, project writers complete the first draft individually with the help of a strategy of the thinking maps, the acquired knowledge (i.e. language and thinking skills) from lectures and their former writing experience. Through strategies of peer review and group discussion, they receive, analyse and evaluate the comments, make a judgement on which suggestions benefit better thinking and language performance, and regulate and revise the first draft, eventually completing the final draft. The collaboration helps project writers bridge from current competence in language and thinking to a higher level, resulting in better writing performance. Unlike teachers' assistance (i.e. lecture) stressing the generic language and thinking demand for completing a writing project, the interaction in strategies of peer review and group discussion helps project writers detect their own mistakes to fill specific

knowledge gaps in different groups. This contextualised assistance serves as scaffolding (ZPD) to situate the new skills' learning in a more authentic environment. In this stage, project writers use these two strategies to receive the "guidance-in-context" from their peers and internalise the new knowledge in the final draft (Tabak and Reiser, 2022, p.56).

On the other hand, peer reviewers could also develop their cognitive competence. When revising the essay, some may acquire new knowledge by reading through it, especially for those who are weak in language and/or thinking. Meanwhile, they detect the mistakes and provide comments by linking to the knowledge acquired from lectures and their writing experience. When exchanging their comments in groups, they, for the second time, may acquire the information they barely knew before and receive instant feedback from the project writer and other peer reviewers. All the messages they obtain through the interaction in revision and discussion are expected to contribute to their cognitive development and fill the knowledge gap they may not realise before. Over time, it is assumed that these reviewers will outperform their current writing performance when becoming project writers. By acquiring and practising more CT knowledge, they are supposed to need less time to complete the writing projects. Their development could offset the drawbacks of the "unfair" time allotment in writing when someone chooses the last unit to have less time to revise and accomplish the project. Briefly, students bridge from their present abilities to a higher level and contribute to a better learning outcome and higher mental development for both project writers and peer reviewers. In addition, this is the central stage, especially for the project writers, to revise the draft mainly with the assistance of peers and complete the final draft.

In the presentation stage, the teacher posts all the writing projects on an electronic teaching platform for students to read after getting students' permission. This platform serves as a material tool for PBL participants to learn about other groups' essays. They could reflect on the topics they are interested in by analysing other essays' perspectives or organisation. In the post-project stage, all the participants complete questionnaires to investigate their attitudes and to provide suggestions for future PBL teaching. When making suggestions, respondents reflect on whether CT instruction promotes or militates thinking development and which strategy or activity is effective.

In addition, reflection activates students' thinking to evaluate their performance in different PBL stages and then detect their weaknesses and strengths, contributing to their self-regulation for better performance. As a result, they could recognise the PBL's pros and cons and provide more precise suggestions to help them overcome weaknesses and maintain advantages. It implies reflection leads to a higher mental function to deeply understand PBL and identify their weaknesses and strengths when joining PBL.

On the other hand, teachers could examine the positive and negative sides of PBL teaching and regulate the following instruction to promote the project work every year based on the feedback (Alan and Stoller, 2005). In this sense, reflection and future regulation may motivate learners to engage more in PBL, resulting in a better learning outcome. Briefly, this section illustrates how CT instruction benefits students' thinking based on the critical concept of mediation and ZPD in SCT.

4.4 Thinking Maps

Serving as a language for thinking, the thinking map strategy helps students and teachers make thinking visible and organised (Hyerle and Alper, 2011). There are eight categories of "circle, bubble, double bubble, tree, brace, flow, multi-flow and bridge" (Grigg and Lewis, 2019, p.111). Different types demonstrate specific thinking processes, like "defining in context, describing qualities, comparing and contrasting, classifying, part-whole relationships, sequencing, cause and effect, and seeing analogies" (ibid., p.111). Teachers' modelling of drawing thinking maps helps students recognise and evaluate the writer's thought process, especially in reasoning (Fisher, 2011). Then, students could apply such knowledge to create their thinking maps for writing projects, which could be helpful for students, especially those who lack experience in critical writing.

Besides, it is a common-used pre-writing activity (Al-Shaer, 2014; Lee, 2013), which requires students to use high-order thinking skills to clarify topics and organise ideas (Daley and Torre, 2010; Liu, 2011; Rahman and Ambreen, 2018; Rosen and Tager, 2014). Moreover, it raises their CT awareness of inferential relationships (Kaeppel, 2021). This thinking process leads to systematical writing essays (Rahman and

Ambreen, 2018). In this sense, it could be employed in thinking training as a pre-writing strategy in this PBL course design.

4.5 Peer Review

As an indispensable step in writing critically, peer review means learners evaluate their peers' written projects (Dong, 2017; Nosich, 2012). Students need to consider the conception made, whether the evidence could support the findings, whether there is any gap in the evidence (Moon, 2008) or whether the reasoning process can be justified. Since students have different understandings of CT skills, they are likely to have different standards to comment on the thinking quality. Therefore, teachers should provide a rubric or checklist with peer reviewers to make a "fair" decision. At the same time, it also benefits the project writers for assessing or reflecting on the thinking quality when writing their essays. Dong (2017) points out the necessity of placing thinking elements, such as logic or coherence, into ESL/EFL writing criteria, especially argumentative writing. In this study, due to the exploratory nature of the effectiveness of CT-oriented PBL and students' inexperience in thinking knowledge in writing, a checklist adopted from the acceptable scale of the rubric for assessing CT skills in writing is provided for students to check whether they properly use thinking skills (Appendix I).

Peer review in this study involves a separate and collective work process: reviewers first correct the essay individually and then share their comments with other peer revisers and project writers. Individual work is the premise of collaborative review in this two-way process. After the first class slot, every reviewer applies the CT knowledge to evaluate the first draft following the thinking checklist. The reviewers and the project writer discuss their suggestions in the second class slot. During the discussion process, reviewers not only exchange their perspectives but also check whether other reviewers' corrections are proper and tackle their confusion. In this sense, project writers do not passively receive feedback and are unsure whose suggestions are more reasonable but interact with the conflicting ideas. They may encounter many difficulties, such as low-quality feedback (Hanjani and Li, 2014; Yu *et al*., 2021) or free riding (Han and Hyland, 2015; Meijer, 2020). However, collective peer review allows students to discuss obstacles and verify conflicted comments (Abdullah, 2018; Villarreal and Gil-Sarratea, 2020) and learn in an "anxiety-free

environment" (Beiki *et al*., 2020, p.19). These benefits are more likely to enhance individual performance and thinking skills to reflect on diverse perspectives (Kölbel and Jentges, 2018; Silva *et al*., 2016). Mujtaba *et al*. (2021) prove that collaborative groups outperformed the individual group in the amount of mistake detection.

4.6 Group Discussion

Group discussion in EFL writing typically refers to meaningful interactions among students in small groups discussing how to plan writing projects before writing individually (Li *et al*., 2020; Li and Zhang, 2021). Specific to this study, it also includes exchanging peer reviewers' feedback on the essays after writing the first draft. As a backup for the strategy of thinking map before writing and peer review after writing, pre-writing group discussions concerns revising thinking maps drawn by project writers for better individual planning of the first draft and post-writing collective interactions exchange peer comments contributing to better performance in the final draft. Focusing on solving problems through thinking maps and peer reviews makes group discussion more targeted with a clear purpose. These structured tasks generate more reflective content in the discussion, especially in aspects of content and organisation (McDonough and Neumann, 2014). In this sense, group discussion would develop students' CT.

4.7 Summary

This chapter first presented the theoretical baseline (SCT) behind a PBL teaching design to teach CT in an EFL course by articulating why PBL could be used and what CT instruction in PBL is to teach CT. Next, it presented the theoretical framework of CT-oriented PBL delving into how CT instruction is embedded with different PBL stages of PBL to maximise the benefits of CT improvement. Finally, it illuminated the possibility of using three language strategies as a CT training tool to increase students' thinking skills and awareness.

PART II METHODOLOGY

In previous chapters, I provided the general context: issues of developing students' CT skills. Then, I reviewed the relevant literature: critical thinking and project-based learning, including their definitions, teaching practice and the connections between CT, PBL and the English language. Furthermore, I developed a new PBL framework focusing on CT development on the basis of SCT. Therefore, this study aimed to explore the feasibility and effectiveness of CT-oriented PBL in the CE course. Accordingly, the research objectives had been set out to identify the extent of CT improvement, investigate students' attitudes and explore how CT strategies affect CT enhancement. Driven by the research aims and objectives, research questions in this study were formulated to concentrate on identifying teaching effectiveness and explore how CT-oriented PBL affected students' thinking improvement. The following three chapters outline the rationale for my research design of using a case study and how it fitted into the research questions (Chapter 5). Data collection and analysis methods and ethical considerations will be discussed next (Chapter 6). Finally, it assesses the reliability and validity of the research design (Chapter 7).

Chapter 5 Methodology Part I

5.1 Introduction

This chapter presents the theoretical baseline of the methodology in this study. It begins with the research paradigms from epistemological and ontological perspectives. The research strategy of the case study is followed by discussing its rationale and criticism. Then, a single-case study design is explained due to the research purpose of exploring the CT development of the same students at two different points, followed by a mixed research method of collecting and analysing data. Finally, it discusses the researcher's role as a teacher and researcher in this study.

5.2 Research Paradigms

A paradigm is "a shared belief system or set of principles the identity of a research community, a way of pursuing knowledge, consensus on what problems are to be investigated and how to investigate them, typical solutions to problems, and an understanding that is more acceptable than its rivals" (Cohen *et al*., 2018, p.8).

Different paradigms relate to varying visions of studying social issues, affecting how researchers understand the world and allowing them to organise their thinking about the research. In social science, Cohen *et al*. (2018, p.6-7) identify four assumptions: ontology, epistemology, human nature and methodology. Ontology concerns "the nature of natural phenomena" (i.e. views of reality). Epistemology mainly concerns "the base of knowledge and how it can be acquired and how communicated to other human beings" (i.e. ways of understanding, knowing about the researching the reality). Human nature concerns "the relationships between human beings and their environment". Several paradigms are developed to understand the world based on the above assumptions, such as positivism, interpretivism, objectivism, constructionism, pragmatism, and realism (Bryman, 2016; Lather, 2004; Lukenchuk, 2013). This study utilised two primary paradigms in social science: positivism and interpretivism. Irshad (2017, p.88) provides a clear outline by presenting their nature, methodology and research methods (Table 5.1).

Table 5.1 A summary of two major paradigms

Paradigm	**Ontology** *(What is reality?)*	**Epistemology** *(How can I know it)*	**Methodology** *(How do I approach finding out?)*	**Methods** *(What techniques are best to use?)*
Positivism	There is a reality "out there"	It may be possible to measure given the right tools	Experimental or survey research	Quantitative method predominates
Interpretivism	Understanding is created in the mind of individuals	Reality needs to be interpreted	Often linked to observation, phenomenology, discourse analysis	Qualitative methods predominate

The Positivist paradigm advocates that social science is the same as natural science, concerning "discovering natural and universal laws regulating and determining individual and social behaviour" (Cohen *et al*., 2018, p.8). The social world can be studied scientifically, focusing on the facts and statistics to explore the causes and consequences (Denscombe, 2017). Therefore, it is mainly related to the quantitative method. In contrast, interpretivist paradigm "respects the differences between people and the objects of the natural sciences" and strives to research the social world by people's subjective interpretations and views (Bryman, 2016, p.26). Therefore, it is mainly relevant to the qualitative method.

This study focuses on exploring the feasibility and applicability of CT-oriented PBL intervention to improve students' CT. First, this exploration is conducted by assessing whether students will develop their CT in assessment after completing the intervention. Specifically, CT development manifests in scores of standardised CT tests and the frequencies of using CT skills in writing projects. It means the objective results exist independently with properties that can be studied statistically (Table 5.1). In this sense, it fits into the positivist paradigm. Second, it is also be conducted by investigating students' views about the PBL effectiveness to improve thinking by questionnaire and exploring their interpretations about their CT development. It means the nature of this exploration is based on students' views and interpretations of their experiences (Table 5.1). In this sense, it also fits into interpretivist paradigm. Therefore, this study adopted a mixed method design of mixing two paradigms of positivism and interpretivism with quantitative and qualitative approaches with multiple data collection.

As previously discussed, these two paradigms are contradicted. Whether they are commensurable or can coexist, or be integrated is open to question (Bergman, 2011; Biesta, 2012; Hammersley, 2013). Cohen *et al*. (2018) claimed the mixed research method argues for the compatibility of different paradigms, or at least, they can coexist and work together to resolve research problems. Creswell and Guetterman (2018) put forward another philosophical paradigm: pragmatism, to provide a possible solution to incompatibility. Namely, "using procedures that 'work' for a particular research problem under study and that you should use many methods when understanding a research problem" (ibid., 2018, p.547). In other words, as long as they successfully answer research questions based on the research purpose, researchers could be eclectic in research designs, data collection and data analysis. Therefore, the quantitative data of measuring students' CT development and the qualitative data of investigating their attitudes and perceptions are equally significant and compatible because it helps gain a comprehensive understanding of the research questions.

5.3 Research Strategy

A case study is an in-depth or detailed investigation of a specific real-life project, policy, institution, program or system from different perspectives to capture a small

sample's complex and unique features (Cohen *et al*., 2018). As a widely-used research method in social science, it helps researchers deeply understand complex learning and teaching processes, the contexts, individual(s) or communities under education settings (Hamilton and Corbett-Whittier, 2013). Empirical findings and interpretations of the case study allow practitioners to explore significant case features and consequently understand, evaluate, and improve their teaching practice. This section introduces the rationale, criticisms of the case study, a single-case study design, diverse research data collection, and my role as a researcher.

5.3.1 Rationale for case study

"The form of research question posed" is the primary and the most crucial factor for researchers to choose (a) research method(s) when conducting research (Yin, 2018, p.9). This research explored the applicability and effectiveness of using a PBL approach to develop students' CT skills in one Chinese higher education setting. A case study research method helped me explore what developed and how the development happened in a specific context. In accordance with the present research purpose, three research questions, including two "What" questions (i.e. what might be the impact of CT instruction on EFL students' CT in higher education and what are their attitudes towards and perceptions of the effectiveness of CT-oriented PBL intervention) and one "How" question (i.e. how CT strategies in PBL might impact the critical thinking improvement), were raised.

"What" questions here were to explore to what extent students' thinking improved after PBL intervention and their attitudes towards PBL and CT strategies. Rich data collected from CT tests and questionnaires were statistically described and interpreted in words. Deep data collected from students' writing projects (i.e. the first and final draft) were deductively coded to provide detailed evidence to present their development of content-specific thinking skills. This result triangulated the CT test in thinking skills. At the same time, the frequency statistics in writing triangulated the CT test in disposition. Both quantitative and qualitative data were to answer the first two questions with exploratory and evaluative nature here. In addition, the "how" question was to explain their dynamic development in the CT-oriented PBL learning process. Rich and in-depth data (i.e. questionnaire, interview) were interpreted to respond to the question with explanatory nature here. As a research method, the case

study design could answer exploratory-oriented "what" questions and explanatory-oriented "how" questions (Yin, 2018). In this sense, the case study was a result of reflecting on research purposes and research questions.

This research aimed to explore a feasible channel to develop thinking in a natural social setting (a Chinese university). Devoting my effort to one instance (the CE course) helped me delve into issues in depth and possibly contribute to localising PBL teaching. Furthermore, this research mainly focused on investigating and interpreting the relationships between PBL and students' CT improvement from a student's perspective. It was necessary to learn how various aspects (i.e. students' attitudes, performance changes in thinking, and their behaviour in completing writing projects) interacted in order to explore whether this relationship existed and how it happened. The case study worked well to view the case as a whole to interpret what was happening from a holistic view instead of dealing with the different strands separately. Finally, multiple data sources were required in this study to answer "what" and "how" research questions. The features here are consistent with almost defining characteristics of a case study.

5.3.2 Criticism of case study

The case study research method allows researchers to focus on the case closely and conduct the research thoroughly utilising multiple data sources and flexible research methods in a holistic manner (Denscombe, 2017; Hamilton and Corbett-Whittier, 2013; Tight, 2017). It was also more feasible to carry out a small-scale study owing to time limitations or resource restrictions within a university setting. However, Mackey and Gass (2005) point out the limitations of a small number of participants, or it merely focuses on some specific cases, resulting in criticisms of the generalisability (Duff, 2012, 2014; Merriam and Tisdell, 2015; Tight, 2017) and objectivity (Duff, 2014; Flyvbjerg, 2006; Robinson, 2013; Saldaña, 2011).

In generalisability, this study particularised findings from the selected case and explored the likelihood of broader relevance to PBL teaching theory rather than universalising quantitative results from enormous random samples. Therefore, generalisability here cannot be evaluated by a common standard of quantitative research. In subjectivity, participants were potentially affected by picking up clues

from me, a lecturer and a researcher in this study, to cater to the research need. It means the power of a teacher-student relationship or the face-saving culture may influence students' opinions (Duff, 2020). Besides, Duff and Anderson (2016) and Hancock and Algozzine (2006) also point out some subconscious bias from students' subjective interpretations due to biased perceptions and recollections of the events when taking part in an interview or questionnaire. In order to reduce their bias, ethical issues were considered when processing data, especially from the interview and the questionnaire (it will be discussed in Chapter 6, section 6.6).

No matter where the bias came from, the present study's credibility would be negatively affected by the research objects (i.e. students) or the researcher myself. Several strategies, for example, triangulation, member reflections, or peer review, were suggested to minimize such risk (it will be discussed in Chapter 7, section 7.2).

5.3.3 A single-case study design

The central purpose of this research is to explore whether CT-oriented PBL intervention improves students' thinking in a CE course. Following on the same participants and observing if their CT performance has changed after delivering an intervention course enables me to build an overview and a deep understanding of the changes. In this sense, a longitudinal case study, featured in investigating processes and changes of the same single case, is optimal to serve this research purpose (Hamilton and Corbett-Whittier, 2013). This study is a single-case study due to the demand for a longitudinal case study to track the same individuals over time. It is consistent with Yin (2018), who claimed the longitudinal case was a rationale for a single-case study design. Specific to this study, a longitudinal case-based approach achieved the primary purpose of this research by tracking the same samples (i.e. groups of first-year undergraduates) to specify their CT transformations at two different points (i.e. before and after PBL) for one academic term following a before-and-after logic.

Without understanding what a case was, the case study cannot be really considered complete (Duff, 2014). Therefore, I need to identify the case at the outset. This process involves two different steps: "defining the case and bounding the case" (Yin,

2018, p.28). The sampling strategy was considered next after identifying the case. The research context was also introduced in this section.

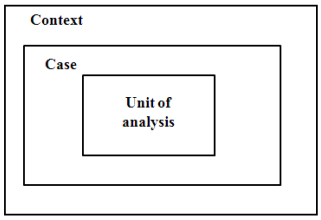
5.3.3.1 Defining the case and its boundaries

The case in a case study design means different units of analysis (Duff, 2012). It is typically defined as "a person or a small number of individuals on their own or in a group" in applied linguistics (Duff, 2014, p.233). Although working in different groups, students are still seen as one unit since this case study was to explore the impact of PBL intervention on overall performance at an exploratory stage instead of investigating factors influencing individual performance. In this light, the case in this study was defined as a group of EFL learners in a Chinese university who participated in PBL learning.

Although this was exploratory research, it was predicted that PBL would be an effective approach to enhance students' CT. It was expected to provide empirical evidence of its effects, thus making it possible to testify whether the CT-oriented PBL was applicable in Chinese higher education. Therefore, the case was bounded in the targeted students who a) were all first-year undergraduates without previous CT or PBL learning experience, b) had comparably better English proficiency on average (i.e. at least 105 out of 150 points, 70% of total mark in Gaokao), c) can regularly take part in the CE course and d) learnt English as a foreign language and had the same L1 (Mandarin). The case boundaries could also be defined as criteria for recruiting samples.

The current study assessed the same group of first-year undergraduates at two different time points in order to investigate their thinking changes over time. The units of analysis were at one level of those who attended the CE course. Therefore, the holistic single-case study was designed without logical subunits (Figure 5.1).

Figure 5.1 A holistic single-case study design



5.3.3.2 Sampling

Convenience sampling was used in this study to choose from my own institute as the participants in this study out of two considerations. On the one hand, it was challenging to recruit a practitioner who could deliver the intervention programme because it was innovative and has not been popularised in China. On the other hand, it was convenient for me to conduct PBL teaching and select targeted students as participants. This sampling method is the most prevalent in L2 studies, and students in researchers' own institutions are prime examples of convenience samples (Dörnyei, 2007). Although convenience is the primary basis for sample selection, it should not be the only criterion. Otherwise, it may reduce the validity of the research to some degree (Tight, 2017). Dörnyei (2007, p.99) points out that "convenience samples are rarely completely convenience-based but are usually partially purposeful". It means participants should possess key features for a research purpose in particular. Specific to this study, participants should have four particular characteristics as mentioned in the preceding section 5.3.3.1.

Firstly, participants did not have any experience in learning English using PBL and CT strategies or had minimum experience. In this light, they were thought to stand at the same starting point in general, which may enhance the validity. In this sense, the first-year undergraduates were the most potential candidates due to their lack of CT training and PBL learning in higher school.

Secondly, samples were supposed to have higher English language proficiency than the average level. As discussed in Chapter 2, section 2.4.1, EFL learners' CT development was contingent to a great extent on their English abilities to express their opinions clearly and logically. If participants with quite low English levels (e.g. lower than the passing score: 90 out of 150 in Gaokao) were selected for this study, they might find it difficult to demonstrate their thinking even if they are developing their thinking. At the initial stage of exploring PBL effectiveness, language proficiency should not be a major barrier to applying CT knowledge. In this study, students whose English scores in Gaokao reached 70% of the total mark (i.e. 150), namely 105, were supposed to be the participants. Specifically, students who majored in stomatology were selected because their average score was 123, which was the highest at this

university. In this sense, they were believed to have the abilities to display their thinking performance and be suitable for the English requirement for the intervention.

Thirdly, the samples were supposed to be able to attend the class regularly and have enough time and energy to complete projects. Finally, students should be EFL learners with the same L1 because PBL teaching was conducted in an EFL course and having the same L1 may help them communicate and cooperate more efficiently.

5.3.3.3 Participants

No ideal and unified sampling selection criteria can be applied to all the case studies for different research purposes in diverse contexts. Following the above criteria, 98 stomatological first-grade students (48 in class one and 50 in class two) in my university were chosen as the samples for this study. The sample size of 98 ensured the statistical sense ($n > 30$) and decreased the risk of affecting data collection and analysis in case someone may withdraw from the research for various reasons. Although the 98 students were in two separate classes, they could be regarded as one group because they studied exactly the same modules and had similar backgrounds (Table 5.2).

Participants were in their first year of study for a five-year degree program in the School of Stomatology, Binchou Medical University in China. Before entering the university, most of them had already learnt English as a foreign language for at least nine years. They were expected to reach level five (the equivalent of IELTS 5.5) when finishing their two semesters of College English in the first academic year. This study was conducted in the first semester from 16th September to 31st December 2019.

Table 5.1 shows the overview of participants' (n=98) demographic information. They are aged 17 to 20. The total number of female (n=54, 55.1%) and male participants (n=44, 44.9%) is almost equal in this medical university. Regarding the age, most of them were 18 and 19 (n=92, 93.9%) when the current study was conducted (Table 5.2).

Table 5.2 Participant Demographics

		PBL class 1 (n=48)		PBL class 2 (n=50)		Total (n=98)	
		N	%	n	%	n	%
Gender	Male	21	43.8	23	46	44	44.9
	Female	27	56.3	27	54	54	55.1
Grade	1st	47	100	47	100	98	100
Age	17	2	4.2	2	4	4	4
	18	25	52.1	29	58	54	55.1
	19	20	41.7	18	36	38	38.8
	20	1	2.1	1	2	2	2

(PBL class 1= Stomatology class 1, PBL class 2= Stomatology class 2)

This study was conducted in the Binchu Medical University context, which serves as a pseudonym for a confidential purpose. It is an average-sized medical university with about 15,382 full-time undergraduate students in China. This university has 16 schools with 34 medical-related majors (e.g. clinical medicine, stomatology, paediatrics and nursing), and stomatology is one of the best-developed majors in this university. It was evaluated excellent level for their undergraduate teaching work last time by the Chinese Ministry of Education. Since 2018, as the first batch of universities taking part in an Undergraduate Teaching Project program conducted by the Ministry of Education, it has encouraged all the teachers to reform teaching for different levels of students. In this sense, it provided a supportive platform for me to use PBL and develop students' thinking in my class.

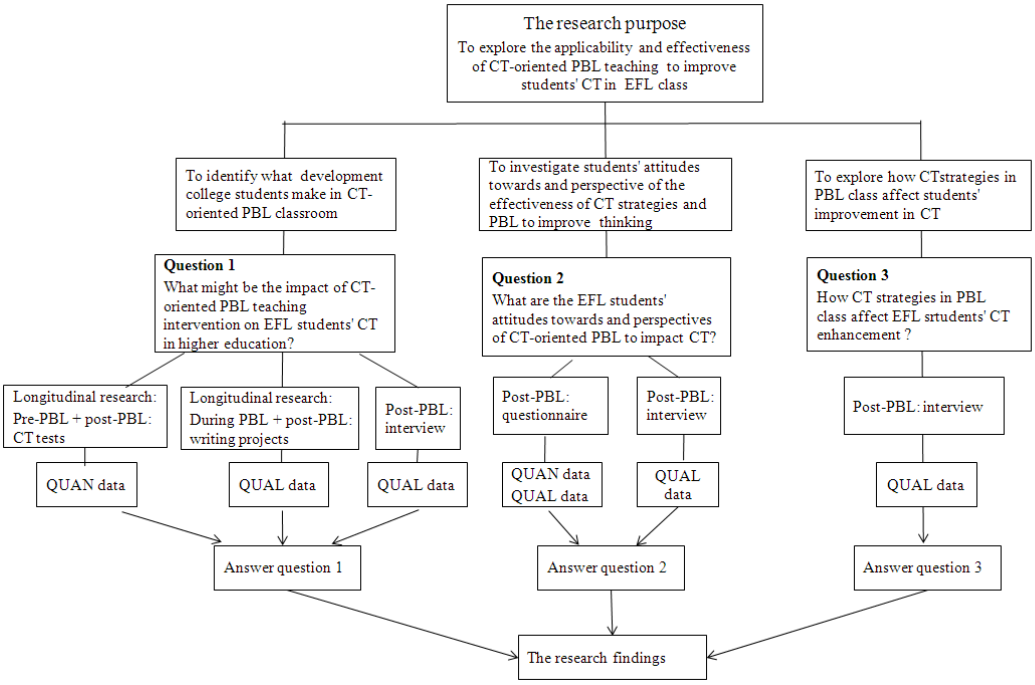
5.3.4.Mixed methods

On the basis of the research questions and case study design in this study, any single data collection method (i.e. a quantitative or qualitative method) was deficient in providing the answers ultimately to achieve the purpose of this study. A mixed research method of collecting both data was consistent with one case study feature that allowed the researcher to collect data from multiple sources. Furthermore, using dissimilar methods to collect and converge data from different sources could triangulate this case study (Cohen *et al*., 2018; Yin, 2018).

The central purpose of this study was to explore whether PBL teaching was applicable to improve students' thinking in general. If it did, the results might be so broad that a follow-up was needed to understand further how it worked in such a learning process.

The rationale for data collection in this study was the combination of quantitative statistical data (i.e. CT tests and questionnaires) and qualitative evidence (i.e. the interview and written texts) (Creswell and Guetterman, 2018). In this sense, the explanatory mixed research design (Figure 5.2) enables a more comprehensive and complete understanding of the feasibility and effectiveness of PBL to be obtained than any single research method. At the same time, it also balances out the limitations of each method.

Figure 5.2 Mixed research data collection within a case study



Collecting data through mixed research enabled me to answer research questions from a holistic perspective to evaluate whether PBL was effective for Chinese first-grade undergradutes to improve their CT. Alternatively, it allowed me to delve into the research questions through their experiences and reflections during this learning adventure. More importantly, it achieved triangulation by multiple data and methods, which could increase the credibility and validity of this research.

5.3.5 Role of the researcher

The researcher performed two roles in this study. In student-centred PBL classes, teachers' roles are coach or guide. Having a dual identity of the teacher and the researcher, I experienced this transformation by myself, which inspired me to

understand PBL teaching and role transition deeply. As a teacher in this PBL class, the role of knowledge supporter manifested in providing the scaffolding for the participants, especially those with less experience in CT and PBL. In addition, double identity gave me a bonus when conducting this research. That means I had more opportunities to directly observe the classrooms and detect potential problems in this learning process. Once students met obstacles, I could help them in time. Meanwhile, I could easily build mutual trust with them. All of these helped with recruiting participants and achieving high response rate in questionnaire. However, this intimate relationship may have a negative side. For example, participants were likely to provide satisfying answers in order to please me or for the sake of "face" when taking part in an interview or the questionnaire. As a result, the results are biased to some degree.

5.4 Summary

This chapter stated the research paradigms from epistemological and ontological assumptions at first. Accordingly, it presented the case study research strategy, including the rationale and criticism and the single-case study design, followed by the data collection and analysis of the mixed research method. Finally, it discussed the role of the researcher.

Chapter 6 Methodology Part II

6.1 Introduction

The previous chapter outlined the philosophical considerations, case study research strategy, single-case study design, data collection methods, and the researcher's role from a theoretical perspective. This chapter aimed to illustrate further how I conducted teaching interventions and collected and analysed data. The ethical parameters of the study were accounted finally.

6.2 CT-oriented PBL teaching interventions

A CT-oriented PBL teaching program was designed based on the framework in the current study (Chapter 4, section 4.3, Figure 4.1). Table 6.1 shows the teaching procedures that each unit will have to follow. Teaching interventions started on 16th September 2019 and lasted for 15 academic weeks. The CE was scheduled to include five units with five topics in the textbook. For instance, love affairs in college, language learning, and generation gaps. Each unit takes three weeks to complete, and students have one 90-minute lesson per week. Students are allocated different roles when working on their projects. The project writer is responsible for writing an essay as the project for researching the topic of the unit. While the other group members act as peer reviewers to be responsible for revising the thinking map and commenting on the essay by group discussion. These peer reviewers are required to prepare for the group discussion to achieve the most significant achievement in groups to discuss the writing errors in class. In other words, they had to revise it in advance before exchanging comments individually. Students took turns acting as project writers and peer reviewers.

Table 6.1 Teaching procedures of CT-oriented PBL teaching

lesson slot		**Teachers**	**Students**	
			Project writers	**Peer reviewers**
1st	before class	-Provide resources (i.e. language) for students to preview	-Preview language points in the provided resources; - Gain a general understanding of the passage	
	in class	-Teach language and thinking knowledge by lecturing the passage in a textbook	-Acquire language and CT knowledge	

	after class		-Do exercises in the textbook; -Note down any questions when doing exercises	
			-Determine a writing title	
2nd	in class	-Respond to questions about students' exercises; -Model how to draw thinking maps; -Provide any help students asked for	-Draw a thinking map	Group discussion: -Revise the thinking map; -Discuss examples or evidence to support the arguments
	after class		-Write the first draft; -Submit it to the teachers online; -Share it with group mates in print	-Review the first draft -Write down all the comments
3rd	in class	-Continue to answer students' questions in exercise - Provide any help students asked for	Group discussion: -Share and discuss the comments on the first draft	
	after class		-Revise the 1st draft and complete the final writing project; -Submit the final one to the teachers online	

6.3 Data Collection: Mixed Methods

A combination of quantitative data (i.e. CT tests and questionnaires) and qualitative data (i.e. interview and writing projects) match the research questions. The contents and purpose of the four data collection instruments (i.e. CT tests, writing projects, questionnaire and interview) are the cruces of this section. Both types of data are separately presented in two sections. For each section, different forms of data are presented based on the demand for answering research questions (Chapter 5, section 5.3.4, Figure 5.2).

6.3.1 Quantitative data collection

In order to answer RQ1 about the impact of PBL on students' CT development, two CT tests were administered before and after PBL intervention: California Critical Thinking Skills Test (CCTST) and California Critical Thinking Disposition Inventory (CCTDI). As the most widely used essential standardised tests of thinking, they aimed to assess students' CT abilities and disposition respectively (Karbalaei, 2012). Both

tests' reliability and validity for university EFL students in different countries have been widely testified (Genç, 2017; Liu and Jin, 2012; Marzban and Barati, 2016; Sadeghi *et al*., 2014). However, designed in English with a western background, both tests still risk not being suitable to assess EFL students from different cultural backgrounds, especially for those with low English proficiency. And this may affect students' CT test performance. Soufi and See (2019) suggest using a translation version or revising some contents in these tests. In this study, the simplified Chinese version of these two CT tests was adopted.

The Insight Assessment company provided individual and group reports of core CT skills and mindset attributes with numerical score ranges and corresponding qualitative descriptions (Appendix C, D). Due to copyright reasons, it did not provide detailed information about item-level data to clients, including answers given by individual test-takers, information about the structure of the instrument (items used to score each variable), and information about correct answers or any other aspect of the scoring algorithm.

The California Critical Thinking Skills Test (CCTST) is an objective test assessing test-takers' critical thinking skills. It is designed in the format of multiple choices with 34 items, providing a statistically comparable score package including an overall score (0-34 points) and three individual core cognitive skills in critical thinking (i.e. Analysis, Inference, Evaluation) (Appendix C1). Students' overall performance was converted into five levels (i.e. not manifested, weak, moderate, strong and superior). At the same time, their performance in three core cognitive skills was tallied into three levels (i.e. not manifested, moderate and strong). Each level has different interpretations of its CT performance (Appendix C2)

The California Critical Thinking Disposition Inventory (CCTDI) is a self-report measurement to reveal participants' attitudes, experiences, and thinking dispositions. It contains 75 six-point Likert-style items in the format of agree/disagree, ranging from "strongly agree" to "strongly disagree". It reports six numerical scale scores and corresponding qualitative ratings (i.e. Truth-seeking, Open-mindedness, Analyticity, Confidence in Reasoning, Inquisitiveness and Maturity of Judgment) to recognise the test-takers' strong and weak areas. These six CT dispositions share the same five score ranges (10-60 points), the same five qualitative categories (from strongly negative to

strongly positive) and the same corresponding qualitative interpretations to evaluate their performance in these six dispositions (Appendix D).

CCTST and CCTDI in this study are indispensable because they provided ample evidence of twofold dimensions in critical thinking of cognitive abilities and mindset attributes (Facione, 1990). Besides, such data collection at the beginning and the end of PBL teaching aimed to testify to whether there was any significant thinking difference. Since English was not a native language for test-takers in this study, their English proficiency may affect their performance in both tests. I chose to use the simplified Chinese version so that students could comprehend the items better and demonstrate their actual attitudes. Both Chinese versions of CCTST and CCTDI have been applied to undergraduates in China (Liu and Jin, 2012; Luo and Yang, 2001, 2002).

In order to answer RQ2 about students' attitudes towards the effectiveness of CT strategies and PBL to improve critical thinking, two questionnaires were used: the Effectiveness of CT Strategies Questionnaire on Improving Students' Critical Thinking (CT strategies questionnaire) (Appendix E1) and the Effectiveness of PBL Questionnaire on Improving Students' Critical Thinking (the PBL questionnaire) (Appendix E2). These questionnaires aimed to investigate students' attitudes towards three CT strategies (i.e. thinking maps, group discussion and peer review) and different PBL components (i.e. teachers' scaffolding role, students' work in group and their awareness of the authentic world).

The CT strategies questionnaire included three sections: a cover sheet, a personal information background section, and students' responses to thinking items. The cover letter described the research purpose and ethical issues. The last section contained three parts. Part I covers students' attitudes towards the effectiveness of CT strategies in improving thinking skills; Part II covers their attitudes in improving thinking dispositions; Part III covers participants' agreement on the effectiveness of three CT strategies separately to enhance their thinking (Table 6.2). The items in Part I and II were designed based on CT components in the Delphi report (Facino, 1990) and the corresponding descriptions of writing requirements in CSE 5 (NEEA, 2018). This questionnaire had 18 items following six-point Likert scales (1= strongly disagree, 2=

disagree, 3= slightly disagree, 4= slightly agree, 5= agree, 6= strongly agree). The mean score was calculated if the sub-skill or the disposition had more than one item.

Table 6.2 CT strategies questionnaire items

The main section	Scales		
Part I	Attitudes towards the development of CT skills	Evaluation	Item 1, 3
		Analysis	Item 2, 5
		Inference	Item 4, 6
Part II	Attitudes towards the development of CT dispositions	Truth-seeking	Item 7
		Analyticity	Item 8, 9, 11, 15
		CT confidence	Item 10
		Open-mindedness	Item 12
		Inquisitiveness	Item 13
		Maturity in judgement	Item 14
Part III	CT strategies	Thinking maps	Item 16
		Group discussion	Item 17
		Peer review	Item 18

The PBL questionnaire included four sections: a cover sheet, a personal information background section, a section relating to students' attitudes towards PBL teaching and a suggestion section. It shared the same structure as the CT strategies questionnaire of Part I and II. Part I in this questionnaire embraced six closed-ended attitudinal questions following six-point Likert scales (1= very disagreement, 6= very agreeable) to investigate their attitudes towards PBL components. Part II contained one open-ended question asking respondents' suggestions for the PBL course.

Table 6.3 PBL questionnaire items

The main section		Items	
Part I: closed-ended questions	Students' general attitudes to PBL components	Teachers' role in PBL teaching	Item 1, 2
		students' working on projects in groups	item 3, 4
		students' awareness of the real world	item 5
	Students' general attitudes to PBL		Item 6
Part II: an open-ended question	Students' suggestions for the PBL course		Item 7

The questionnaire was administered in Chinese, which is the language that participants are more comfortable. It is expected to get accurate and reliable answers.

6.3.2 Qualitative data collection

In order to answer RQ1, data collected from students' first and final writing projects aimed to compare and assess PBL's impact on thinking skills development. Students submitted 153 writing projects, including 98 first and 55 final drafts. Data collection aimed to assess whether their thinking skills developed by comparing their first and last writing. Thus, the writing data would be deleted if participants merely submitted the first or the final edition. As a result, 110 pieces of written texts were valid in this study.

In order to answer RQ1, RQ2 and RQ3, a semi-structured individual interview was considered to be the most suitable method for the following reasons. Firstly, a semi-structured interview helped me to keep the conversation on track and adjust the sequence of interview questions if necessary. The open-ended format encouraged interviewees to put forward their own ideas, perspectives or reactions, allowing me to obtain rich and full information about the topic, including new or unexpected ones. At the same time, their diverse answers stimulated me to ask appropriate follow-up questions. As a result, I could acquire and extend my understanding of the interview questions. Secondly, the individual interview could avoid awkwardness between the interviewees. Since they worked in groups, some may be hesitant to publicly express their true feelings or experiences when stating the negative side, such as dissatisfaction, failure or difficulties.

The prepared questions were divided into four parts with nine questions (Appendix F). The first section (questions 1-3) concerned students' former English learning experiences. The second (questions 4 and 9) concerned their standards to evaluate the writing quality before and after PBL intervention. The third (questions 5-7) was about how each CT strategy impacted interviewees' CT learning. The last category (question 8) was what was the most useful and favourite strategy for interviewees.

One drawback of the interview was the possibility of misinterpreting or misunderstanding participants' responses (Maxwell, 2005). I transcribed each interview into a separate document and sent them to each student, checking whether

there was a correspondence between my preliminary analysis and their perspectives. Everyone returned their response in writing by e-mail, and I revised the scripts accordingly.

In addition, data from an open question in the PBL questionnaire was used to answer RQ2 by investigating respondents' suggestions about this PBL intervention. As a new teaching method emphasising collective, student-centred and CT-oriented learning, understanding how participants respond to it and their suggestions for future course development will be valuable. Although they may leave short or even blank responses, researchers still could receive unexpected responses (Dörnyei, 2009) due to their diverse learning experiences and performance in this course. Researchers could identify any comments students might have beyond answers to the closed-ended questions (Creswell and Guetterman, 2018). In addition, the more suggestions collected, the more comprehensive picture I could build up. At the same time, data from the interview may support their suggestions. Therefore, an open-ended question asking for their advice for future teaching allows students to express their expectations for future learning or for practitioners to reflect and improve their teaching.

6.4 Data Collection Procedures

Having discussed the specific mixed research method to collect data, this section accounts for the results of my pilot study and data collection procedures.

6.4.1. Pilot study

Data collection instruments should always be tried before the actual data collection to ensure it is suitable for the target subjects (Dörnyei, 2009). Pilot studies in this research were CT tests (i.e. CCTST and CCTDI), questionnaires and interviews. The CCTST and CCTDI were tested to determine whether test-takers could finish them within the given time or understand the items. The questionnaires were piloted to check whether respondents could understand the items and whether the items were reliable and valid. Semi-structured interviews were piloted to check whether students could answer the questions and their feelings when being interviewed.

6.4.1.1. CCTST and CCTDI

The pilot study of CT tests was conducted with a group of students who studied phylaxiology (n=47) at the same university on the morning of 3rd September 2019. The User Manual and Resource Guide of CCTST and CCTDI (2020) suggested 45-50 minutes to complete CCTST and 20-30 minutes to complete CCTDI respectively. All the students completed the CCTST within 45 minutes and the CCTDI within 20 minutes, so they had enough time to complete both tests. Before the main study, students also tried the sample questions provided by the Insight Assessment company to help test-takers to be familiar with the format.

6.4.1.2 Post-intervention questionnaires

The purpose of piloting the questionnaire is to check whether the wording was ambiguous, whether the information measured something relevant to the research question, whether the instructions were clear, whether the items were too difficult for the students to respond to, or whether the English version corresponded to the Chinese version. The pilot study was conducted both for my colleagues and students.

Two colleagues with experience in designing questionnaires were invited to part in piloting on the 2nd of October 2019. Regarding the PBL questionnaire, their feedback was mainly about item 7 (i.e. I think PBL learning in writing inspires me to concentrate more on the real world around me, especially on the dialectic point of view). They felt the wording was ambiguous. On the one hand, it can be interpreted as whether PBL could encourage students to pay attention to the world around them. On the other hand, it can be understood whether PBL helped them learn about the topic they were interested in a critical way. Moreover, item 9 (i.e. I think PBL improves my critical thinking) had already expressed a similar meaning. It was suggested to delete item 7. In terms of the CT strategies questionnaire, they suggested that item 6 (i.e. group discussion helped me collect evidence from different sources in support of an argument) should be changed from "group discussion" to "group discussion for thinking maps". Since group discussion occurred in two steps of revising thinking maps and writing projects, this change makes more specific to the step of revising thinking maps. Otherwise, the respondents may be confused about which step this item intends to ask.

After revising the questionnaire according to colleagues' suggestions, thirty students who were the participants in this study were invited to take part in the pilot study on 8th October 2019. Because the questionnaires aimed to investigate students' attitudes and the items are closely related to PBL learning, it is possibly difficult for students without relevant learning experience to understand the contents of the questionnaire and express their attitudes. In this light, the participants in this study were supposed to be the optimal candidates. They spent about 5-7 minutes completing the PBL questionnaire and about 10-15 minutes completing the CT strategies questionnaire. After finishing all the questionnaires, they did not give any negative feedback. I also asked whether they understood the difference between "slightly disagreeable" and "slightly agreeable". They explained their understanding by saying, "the extent is limited to my disagreement or agreement".

6.4.1.3 Semi-structured interview

This pilot study was conducted on the 23rd of December 2019. Three students volunteered to take part in the interview. The piloting results showed that they understood the interview questions and were not offended when being asked all the interview questions.

6.4.2 Data collection procedures and results

The data collection process was introduced in this section. The data were collected from 2nd September to 27th December 2019 in two classes, in three stages: before, during and after PBL teaching intervention (Table 6.4). Before and after PBL, the CCTST and CCTDI data were collected from all the participants. During and after PBL, students' first and final drafts were collected from the same students. After PBL, 17 students volunteered to be interviewed, and all the participants were invited to complete PBL and CT questionnaires.

Table 6.4 Overview of data collection procedures

RQs	Instruments	Data collected from	Pre-PBL	During PBL	Post-PBL
RQ 1	CCTST + CCTDI	Two PBL classes	√		√
RQ 1	Writing projects	Two PBL classes		√	√

RQ 2	PBL + CT Strategies Questionnaire	Two PBL classes			√
RQ 1+2+3	Interview	17 students from two PBL classes			√

6.4.2.1 CT tests: CCTST and CCTDI

Students completed the two tests before starting their PBL intervention at the beginning of the semester and then after PBL at the end of the semester. For the pre-PBL CT tests, both PBL classes completed the CCTST in 50 minutes and CCTDI in 30 minutes before the first class on the same day. The post-PBL CT tests were conducted separately after the last scheduled CE course of the semester (Table 6.5, Table 6.6). Before taking part in the tests, participants were informed that the results were kept confidential only for research purposes and not the standard to assess their academic performance.

Table 6.5 The statistical overview of matching data on CCTST

PBL Class	Students number	Number of valid test/ Total number of valid test (pre-PBL)	Number of valid test/ Total number of valid test (post-PBL)	Number of valid test	Pre-PBL	Post-PBL
1	48	47/48	47/48	47	09/16/2019	12/30/2019
2	50	49/50	48/50	48	09/16/2019	12/31/2019

Table 6.6 The statistical overview of matching data on CCTDI

PBL Class	Students number	Number of valid test/ Total number of valid test (pre-PBL)	Number of valid test/ Total number of valid test (post-PBL)	Number of valid test	Pre-PBL	Post-PBL
1	48	48/48	46/48	45	09/16/2019	12/30/2019
2	50	48/50	48/50	48	09/16/2019	12/31/2019

Table 6.5 and Table 6.6 display the same response rate of 98% in both CCTST and CCTDI before PBL intervention; the response rate is 97% and 96% respectively after PBL. "Only tests with at least 60% of the questions answered and with at least 15 minutes on CCTST were regarded as the valid results", according to *CCTST: 2020 User Manual and Resource Guide* (Insight Assessment, 2020a, p.11). Likewise, the CCTDI tests were valid with responses to "at least 60% of the items and at least 5 minutes", according to *CCTDI: 2021 User Manual and Resource Guide* (Insight

Assessment, 2021, p.10). The CCTST data was only collected from the test takers who took part in this test twice before and after PBL intervention, aiming to investigate their thinking development. The same as the CCTDI. Therefore, the valid number of CCTST was 95, and the CCTDI was 93.

The results showed high response rates because students were very excited and eager to take part in the tests. Almost all the participants were curious about their CT levels since they were never tested. After PBL, I suppose they wanted to learn about whether their thinking level improved raise or not. Their curiosity could partly explain why the response rates were comparatively high (above 95%).

6.4.2.2 Post-intervention questionnaire: PBL and CT strategies questionnaire

Both questionnaires were conducted at the end of the semester on the 3rd and 4th of December 2019. Students volunteered to complete the PBL questionnaire in 10 minutes and the CT strategies questionnaire in 20 minutes. Before conducting the questionnaires, the same instructions were given as CT tests to ensure students knew their rights and responsibilities. Table 6.7 and Table 6.8 present the same response rate of 96% for the CT-strategies and PBL questionnaire.

Table 6.7 The statistical overview of matching data on the CT-strategies Questionnaire

PBL Class	Students Number	Number of valid questionnaires	Data collected time
1	48	47	12/03/2019
2	50	47	12/03/2019

Table 6.8 The statistical overview of matching data on Part I of PBL Questionnaire

PBL Class	Students Number	Number of valid questionnaires	Data collected time
1	48	47	12/03/2019
2	50	47	12/03/2019

Regarding data of the open-ended question in the PBL questionnaire, which asked students' suggestions for future PBL teaching, Table 6.9 illustrates the valid rate is 70% (88, out of 125). Only responses about suggestions for PBL teaching in the CE course are valid. Other responses, like a blank response, comments on the PBL intervention or suggestions for other courses, are invalid.

Table 6.9 The statistical overview of matching data on Part II of PBL Questionnaire

PBL Class	The number of suggestions	Number of valid suggestions	Data collected time
1	69	55	12/03/2019
2	56	33	12/03/2019

6.4.2.3 Post-intervention interview

Seventeen students were interviewed at the end of the semester at different slots from the 24th to 28th of December 2019. The interviews were conducted in Mandarin, allowing students to express their learning experiences and perspectives. After being permitted, the interviews were held in my office with two different audio recorders on the desk. I explained the purpose and process of the interview, and the interviewees finished the consent letter. Additionally, I emphasised that our interview would be audio-taped, and the data would be kept confidential and only for research purposes. Each interview took about 30- 40minutes.

6.4.2.4 Writing projects

During the PBL teaching stage, the first draft writing projects (n=98) were collected at the end of week 2, 5, 8, 11 and 14. After PBL, final drafts (n=55) were collected. To explore whether students improve their thinking skills, I need to compare their first and final drafts; therefore, it is valid only if they submit the first and the final drafts. In this study, 55 students submitted both drafts; therefore, the valid number is 110.

6.5 Data Analysis

This section discusses different analysing data methods (Table 6.10). Firstly, a paired sample t-test or a Wilcoxon's signed rank test was used to assess whether a statistically significant difference existed. Secondly, content analysis was chosen to analyse students' writing projects. Written texts were coded into three categories, and their frequency statistics before and after PBL were provided. A Wilcoxon matched-paired signed rank test computes the difference between each set of matched skill pairs before and after PBL class. Thirdly, a repeated measures ANOVA or a Friedman test was used to evaluate whether there were significant differences among three CT strategies and three PBL components. If there was an apparent difference, a separate Wilcoxon signed rank test was to assess which strategy or PBL component is

believed to be the most or the least effective. Finally, content analysis was used to analyse the interview data to understand the interviewees' performance and feelings.

Table 6.10 Methods for data analysis

Instruments	Data Analysis
CCTST and CCTDI	SPSS: Descriptive analysis, a Paired samples t-test or Wilcoxon's signed rank test
Students' writing	Content analysis: Nvivo: Categories coding, frequency of occurrence; SPSS: Wilcoxon's signed rank test
Questionnaire	SPSS: Descriptive analysis, a repeated measures ANOVA analysis or a Friedman test, a separate Wilcoxon signed rank test. Content analysis: Nvivo: Categories coding, frequency of occurrence;
Interview	Content analysis: Nvivo: Thematic coding, frequency of occurrence.

6.5.1 CT tests: CCTST and CCTDI

A paired samples t-test (assuming assumptions for the test are met) would discover whether there existed statistical differences by measuring the same group of participants at two points in time (pre-and-post-PBL CT tests) (Cohen *et al*., 2018; Dörnyei, 2007). For example, whether participants improved CT skills could be assessed by checking whether they significantly increased scale scores of the CCTST after teaching intervention. The same as students' development in the CCTDI. The CCTST provided an overall score. In contrast, the CCTDI did not because a single strong overall score may potentially hide the weakness of one or more mindset attribute(s) (Insight Assessment, 2021). Therefore, six scale scores of six dimensions were provided as a result.

One limitation of using a t-test is that results can only reveal a statistically significant difference between two variables and whether this result occurred by chance but fails to disclose the strength of a difference. In order to cover this weakness, an effect size can be used to indicate how different the two variables are in this study. Qin and Bi (2015) state that if the effect size is small, the result tends to be insignificant even though it is statistically significant. Therefore, a paired samples t-test revealed

whether there was a significant difference in CT and an effect size indicated how big the difference was.

6.5.2 Post-intervention questionnaire: PBL and CT strategies questionnaire

6.5.2.1 Quantitative data analysis

Descriptive statistics were to calculate the item means and summarise students' tendency of attitudes and perceptions of the effectiveness of CT strategies and PBL teaching. As mentioned before (section 6.3.1), the respondents who chose the value of 1 to 3 indicated their negative attitudes in descending order of the degrees. The lower value was, the more negative attitude they had. While 4 to 6 displayed their generally positive attitude in different degrees with ascending order. Overall, if the mean value was higher than 4.0, the respondents were thought to adopt positive attitudes towards the effectiveness of PBL teaching and CT strategies. In other words, students believed that CT-oriented PBL was helpful for them to improve their thinking in this study. The higher the value was, the more positive their attitude and the higher degree of recognition they granted. Therefore, the mean value above 4.0 was an indicator to show students' recognition of the effectiveness, which was consistent with the research aim of exploring if the pedagogy of CT-oriented PBL was effective or not.

Besides, a repeated measures ANOVA was used to evaluate whether respondents' attitudes towards the effectiveness of three CT strategies were significantly different if the data were normally distributed. If not, a Friedman test was used. This analytical method also applies to assess the effectiveness of three PBL components. Once there was a significant difference, a post-hoc test (i.e. separate Wilcoxon signed rank tests) was employed to measure where the difference occurred. In other words, which CT strategy or PBL component was more effective than others.

6.5.2.2 Qualitative data analysis

A content analysis was conducted to analyse the qualitative data from an open-ended question in the PBL questionnaire. Denscombe (2017, p.313) provided a procedure of content analysis with the following six steps:

Step 1: Choose an appropriate sample of texts or images.
Step 2: Break the texts down into smaller component units.

Step 3: Develop relevant categories for analysing the data.
Step 4: Code the units in line with categories.
Step 5: Count the frequency with which these units occur.
Step 6: Analyse the text in terms of the frequency of the units.

A coding process of content analysis of the open-ended question was as the following steps:

Step 1: Select valid samples of students' responses. As mentioned in section 6.4.2.2, 88 suggestions were collected. This sample size met the requirement of at least 30 for a valid statistical analysis (Cohen *et al*., 2018).

Step 2: Translate the Chinese suggestions into English and ask two colleagues fluent in both languages to assess whether the translation versions reflected the respondents' meaning.

Step 3: Input all the translated suggestions in Nvivo.

Step 4: Develop codes according to students' responses.

Step 5: Calculate the frequencies of codes.

Step 6: Summarise the codes and conclude their suggestions.

6.5.3 Writing projects

Five main methods, including content analysis, grounded theory, discourse analysis, conversation analysis and narrative analysis, are universally employed to analyse talk and text (Denscombe, 2017). Among them, content analysis has an attraction to use both quantitative and qualitative analysis by combing the frequency of occurrence and literal interpretations of the results (Cohen *et al*., 2018). In this study, the qualitative data analysis embraced coding the data according to different thinking skills in writing and calculating the frequency of such codes by the Nvivo. Quantitative data analysis through a Wilcoxon's signed rank test to discover whether a statistical difference exists for the codes of thinking in writing before and after PBL. For example, whether the frequency of the Analysis code significantly increased post-PBL. If it did, it means students improved their Analysis skills in writing statistically. In addition, if the total frequency of codes increased considerably, it was assumed that students raised thinking dispositions in writing. Meanwhile, examples excerpted from the writing data elaborated and supported the above writing

quantitative results. This quantitative result of thinking skills in writing also triangulated quantitative results of the CCTST in generic CT skills.

6.5.3.1 Content analysis process

Based on the suggested coding process of Denscombe (2017, p.313) (section 6.5.2.2), the coding process of content analysis in this study was revised as follows:

Step 1: Choose appropriate samples of the writing projects. As mentioned in section 6.4.2.4, students were required to submit their first and final drafts. Only students who submitted both drafts are valid in order to compare their changes pre- and post-PBL. As a result, 110 writing samples were selected to code, including 55 pieces for first and final drafts sharing the same author. This sample size achieved the requirement of at least 30 that was valid for statistical analysis (Cohen *et al*., 2018).

Step 2: Develop relevant categories and codes in writing projects. This study aims to assess the development of three core thinking skills (i.e. Analysis, Inference and Evaluation) in accordance with the results of the CCTST (simplified-Chinese version). Therefore, these skills were the analytical units of three categories. In addition, ten codes of three categories (i.e. analysis, inference and evaluation) were developed based on the thinking descriptions in writing (Table 2.1). More detailed information will be discussed in the next section (section 6.5.3.2)

Step 3: Develop a coding standard and code the units. Different coders may have different standards for coding writing data. If there was no consistent standard, they might generate different outcomes. Therefore, a consistent standard was needed to be developed, which will be discussed in the next section (section 6.5.3.2).

Step 4: Input all the writing samples in Nvivo.

Step 5: Code the sample writing and combine different codes into three categories separately through Nvivo.

Step 6: Calculate the frequencies of each code and category before and after PBL utilising Nvivo.

Step 7: Analyse the change of such frequencies in order to assess which specific code or category was significantly developed using SPSS.

Step 8: Provide and elaborate literal evidence in writing on the basis of the above quantitative results.

6.5.3.2 A standard for coding CT skills in writing

As mentioned in the last section, it was essential to develop a common standard for different coders to analyse the sample writing since it enhanced the possibility of generating reliable analytical results. This common standard originated from a brief rubric for assessing CT skills in writing based on the descriptions of ten codes with acceptable and weak scales (Appendix I). The acceptable rating means students could use CT skills appropriately to demonstrate logical and coherent relationships between sentences or paragraphs. While the weak rating fails to present such connections. The rubric with two ratings evaluating whether they appropriately use CT skills was in line with the research purpose: exploring whether students improved their thinking skills after PBL. On the other hand, it was more feasible than three or more ratings (i.e. strong, acceptable, unacceptable and significantly weak) for students without experience applying thinking to writing. In this sense, this rubric with an acceptable rating could be a checklist for students to evaluate thinking skills in their essays.

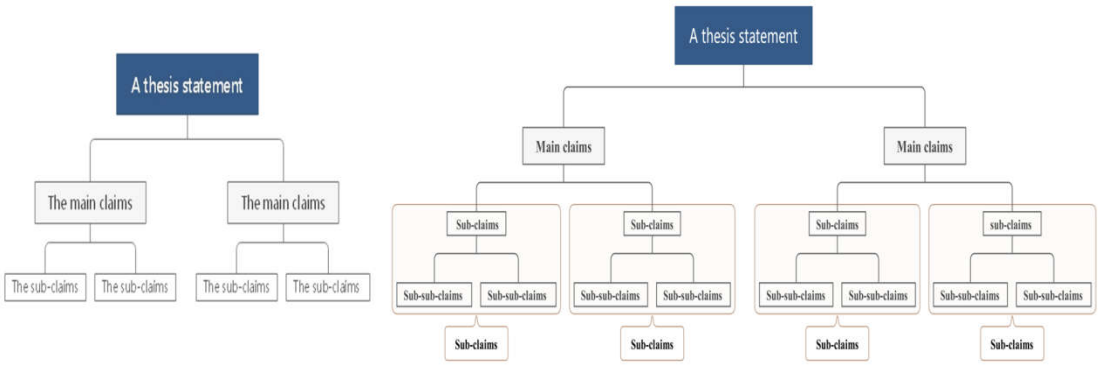
At the same time, coders could use the acceptable rating as a standard to code the writing data. When the frequency of using acceptable CT skills developed, it was assumed that students improved CT. In this sense, even if they use thinking skills in writing but fail to demonstrate logical or coherent relationships, these unit components could be rated weak, which will not be coded. For example, students may use the wrong cohesive device to show progressive relationships. Or the reasons students provided failed to display the causality. In addition, the ratings mainly focused on thinking performance in writing rather than assessing an overall writing performance involving linguistic performance for the research purpose. Therefore, even with language mistakes, such units could be coded as long as the coders understood what students expressed. For example, "At first, it could incentive learning because they may pay more attention to each other only for a better future" (S01-1). In this example, the central part of the sentence "At first, it could incentive learning" could be coded as "providing sub-claims". This sentence could be understood, although that student misused the noun of incentive as a verb. On the contrary, the sentence "because they may pay more for each other only for a better future" (S03-1) failed to be understood, so it could not be coded.

Category 1: Analysis

The category of Analysis was categorised into four codes (1-4) (Appendix I): elaborate thesis statements, provide main claims, provide sub-claims, and compare and contrast.

The 1st code of "elaborating thesis statements" refers to explicitly stating their plans to discuss or prove in an essay. The claims here were divided into the 2nd code of main claims and the 3rd code of sub-claims. The main claims supported the thesis statement, and the sub-claims supported the main claims. These three elements formed a basic writing structure in analysis, which most students followed in this study (Figure 6.1). Sometimes students may use more detailed information to explain sub-claims. In this sense, such sub-claims could be categorised as sub-sub-claims that aimed to support sub-claims (Figure 6.1). No matter what types of sub-claims students used, they reflected students' analytical ability. Therefore, both types can be coded as "providing sub-claims".

Figure 6.1 Two forms of the general writing structure in Analysis skill



An example of analysis coding was: "This leads to the two generations having different ideas. First of all, the two generations have different ideas about marriage and work. The old generation thinks... However, the young generation... Secondly, the two generations have different interests in daily lives...Thirdly, they have different views on setting goals" (S06-1). In this case, the sentence "This leads to the two generations having different ideas" was coded as the main claim, and the sentences "First of all, ... Secondly,... Thirdly,..." were coded as three sub-claims. However, students may write a vague statement to illustrate the main claim although providing the sub-claims.

Another example is, "Various cram schools have opened up their horizons and squeezed their childhood. So is this the right thing to do? First of all, it is undeniable that extracurricular interest classes really enrich the children's vision... Secondly, extracurricular interest classes help children get in touch with the outside world..." (S21-1). In this example, the writer did not provide a clear topic sentence to give the main claim, like "the extracurricular interest classes may have some advantages." Instead, he/she directly wrote two advantages. In this situation, no sentence could be coded as the main claim, but two sentences that stated the advantages (i.e. First of all ... and Secondly, ...) could be coded as two sub-claims.

It was worth noting that transitional signals (e.g. first, second; however, in addition) or the conclusion signals (e.g. to sum up, in brief) were usually used as cohesive devices or markers to connect different claims reminding the coders to focus on the sentences. These words or expressions were introduced in the following section of the Inference category.

The 4th code was "comparison and contrast". Students preferred this critical writing skill, especially when examining how things were similar (comparison) or differed (contrast). This writing skill was introduced in this study in two ways: subject-by-subject and point-by-point. Students were free to use any type they liked. A point-by-point example was: "the old people like chat with family and friends to enjoy simple happiness. In contrast, young people like to chat online and enjoy lonely" (S02-1). S02 used explicit transitional signs to contrast different lifestyles between the older and younger generations. However, some students may not use clear transitional signs to explicit express the relationship between ideas. An example without transitional signs was "The old generation grew (in) poverty, war and disaster. The young generation grew in a time of peace and affluence" (S06-1). Although transitional signs were not used, these sentences implicitly demonstrate the writer has contrasted the two generations by saying they grew up in different environments. In this sense, these sample sentences could be coded.

Category 2: Inference

The inference category was categorised into three codes (5-7) (Appendix I): collect evidence from different sources supporting the claims, use cohesive devices (i.e.

comparison and contrast, progression, transition, and cause and effect) to demonstrate and construct logical and coherent relationships, and draw logical conclusions.

The 5th code was "collecting evidence from different sources". It meant students intended to use multiple information from various sources to interpret, analyse, or evaluate claims. Stapleton (2001) suggested four sources of evidence linked to critical thinking, and three of them could be recognised in writing: personal experience, statistics, and citing sayings. "Citing saying" in this study referred to proverbs or quotations from authorities, traditional culture or official documents. Examples cited from students' writing in the current study were as follows:

-Personal experience: "As for me, one day when I took a trip to the suburb with my English teacher and classmates, a couple of foreigners did a favor (favour) of us. I was about to ask why (they) came to China for a trip when I was shocking (shocked) to find that I could not use my words to form a sentence! Startled and embarrassed, I just smiled to (at) the foreign couple" (S11-1).

-Statistics: "A study of harvard (Harvard) students from 2008 to 2016 found that the number of history majors dropped from 231 to 136" (S24-2).

-Citing sayings: "As an old saying goes: a boat sailing against the current must forge ahead, or it will be driven back" (S05-1).

The 6th code was "using cohesive devices to demonstrate and construct logical and coherent relationships". Students were required to use transitional signs to establish relationships among sentences or paragraphs, such as comparison and contrast, progression, transition, and causality. Examples cited from the writing are as follows:

-Cohesive devices of comparison and contrast: "**contrary to** what is widely accepted, ..." (S05-1) or ".... **Similarly**, parents can't force us to live in their way" (S07-1).

-Cohesive devices of progression: "**What's more**, schools should also improve the atmosphere...." (S16-1) or "**Moreover**, it is a little difficult for them to learn grammar well..." (S11-1).

-Cohesive devices of transition: "**However**, what worried us is that some parents teach their children..." (S08-1) or "**But** some people disapprove this view." (S14-1).

-Cohesive devices of causality: "**Due to** the different ages and experience, ..." (S07-1) or "**Therefore**, everyone should receive education seriously" (S17-1).

The 7th code was "drawing logical conclusions", referring to writers should present their positions following a reasoning process to convince readers to accept the judgment. A summative conclusion was not coded if students simply collected previous information into a shorter overall summary. Because it was a summary rather than a judgement, it deviated from the CT core elements of making a judgment (Cottrell, 2017). Two cited examples showed the difference between a logical and a summary conclusion. A sample of a logical conclusion was, "In a word, it is good for university students to fall in love, but we must have a correct concept of love..." (S27-1). While a sample of a summative conclusion was, "Generally speaking, the humanities not only contribute to educating students but also play an important role in the development of our whole human society" (S24-1). For the first example, the writer made a judgement according to the previous descriptions of the pros and cons of love affairs in college. At the same time, the writer further reminded young people to have a clear and correct concept of love. While for the second example, the writer just briefly summarised essential claims of the benefits and the importance of the humanities without making any judgement. Therefore, the second sample was not the target unit component.

Category 3: Evaluation

The Evaluation category was divided into three types: accessing whether the claims or the evidence are workable, valuable or convincing, discussing the implications and determining what needs doing in a given situation (Appendix I).

The 8th code was "evaluating whether the claims or the evidence are convincing, valuable or workable". It referred to students estimating how reasonable, valuable or practical the claims or evidence were. An example was, "This will make it difficult for them to make progress in the present age when education is increasingly important" (S08-2). S08 evaluated the claim's workable for the current education policy in this example.

The 9th code was "providing implications", suggesting that students should discuss some behaviour or statements' possible effects or results. For example, "..., which

makes us more self-confident when meeting foreigners" (S11-1) or "this has serious effects on the physical and mental health of children" (S08-1).

The 10th code was "determining what needs doing in a given situation". It meant giving suggestions, advice or recommendations. To be able to provide suggestions, advice or recommendations are basic language functions. For example, "The old generation should try to understand rather than blame and respect their views on marriage and work" (S06-1).

In general, the acceptable rating standard in this rubric (Appendix I) was the essential requirement of relevance or accuracy, which can be a checklist for students to peer review. And these unit components could be coded as long as the ten items demonstrate logical and coherent relationships among the sentences or paragraphs.

6.5.3.3 Calculating the frequency of using CT skills in writing

As mentioned in the previous section, the frequency of using core thinking skills was calculated by identifying three categories: Analysis, Evaluation and Inference before and after PBL. The Wilcoxon signed rank test analysed the results to explore whether students increased using CT skills in writing significantly. If so, it showed that students improved their thinking.

The initial coding process of writing data was completed by myself for the drafts. Then one of my colleagues was invited to check the contents of the codes and gave me some suggestions. The codes' descriptions and contents were revised based on her advice and students' actual writing performance. After that, we coded five essays, discussed the divergent results and made a consensus. Next, we coded the first and the final writing drafts (n= 110) separately by Nvivo. The percentage agreement of three themes (i.e. Analysis, Evaluation and Inference) in each piece of writing was displayed. The inter-rater reliability rates of agreement between us for coding Analysis, Evaluation and Inference were 90.58%, 92.97 % and 94.42%, respectively. All of the percentage agreements were above 90% and the rater agreements reached the requirement of inter-rater reliability.

6.5.4 Interview

Interviews in this study aimed to explore students' transformation of a generic CT disposition after PBL, perspectives of CT strategies' effectiveness, and how these strategies affected thinking development. Therefore, the interviews were exploratory with an analytic focus on gaining deeper meaning hidden in the data. More specifically, the data involved students' criteria change for evaluating writing, their understanding of each CT strategy, what specific thinking skills they developed, and what they were concerned about when using different CT strategies. Attempting to satisfy the analytical demand, I utilised the same content analysis method to code, grow ideas, interpret the data and draw conclusions. Meanwhile, this method could provide a quantitative description of the qualitative texts (i.e. frequencies) (Brinkmann and Kvale, 2019; Saldaña, 2016). In other words, it quantified how often specific categories were displayed in texts and then inductively outlined students' critical perspectives from their diverse responses to the same interview questions. For example, in response to a question about what development they made in writing by using thinking maps, interviewees' responses (data) could be coded and then the codes could be subsumed into different categories. Calculating frequencies of those categories made it more likely to draw a statistical conclusion of which aspect students mainly developed in their writing as robust evidence for attesting to their enhancement of thinking skills. Therefore, content analysis was appropriate, taking account of the explorative nature and the data analysis requirement of the interviews in the current study.

6.5.4.1 A brief start list for coding interviews' scripts

Saldaña (2016, p.168) suggests designing "a predetermined start list of codes" based on the interview questions to remind researchers to recognise the contents that were potentially related to the interview questions. The list's provisional codes or anticipated categories were comparatively flexible because they could be revised or added during the coding process (Saldaña, 2016). This coding method was convenient for researchers, esp. for novices, on account of providing an overall and easily-flowed direction to code and generating more specific or accurate codes and categories in the coding process.

As mentioned in section 6.3.2, there were four aspects of open interview questions. The first (questions 1-3) was about students' previous English learning experiences, aimed to identify potential factors impacting students' CT learning, and answer part of RQ3 (i.e. How CT strategies impact CT). The second (questions 4 and 9) was about their standards for evaluating essays before and after PBL intervention, aimed to compare the change of standards, answer part of the RQ1 (i.e. CT learning outcomes in disposition) and validate results from students' writing (i.e. increased awareness to use CT skills). Therefore, the data involved two categories: (1) standards before PBL and (2) standards after PBL. In terms of the third one (questions 5-7) concerned how each CT strategy impacted students' thinking and aimed to answer the RQ3. Based on students' statements, the skills they developed using each strategy were analysed. Besides, it was also necessary to explore what students were concerned about since this was the first time most of them had used such strategies. In this sense, the data involved two categories respectively: (1) the specific writing development students made following each CT strategy, and (2) the issues they were concerned about (i.e. challenges or problems) when taking part in individual strategy. In terms of the last aspect (question 8), it aimed to investigate what was students' most helpful and the favourite strategy and the reasons, answer the RQ2 (i.e. students' attitudes) and validate the results from the questionnaire. An initial coding list provided a brief guideline for the aforementioned anticipated categories (Appendix G1). Categories and more codes were revised or added based on interviewees' responses during the analytical process (Appendix G2).

6.5.4.2 The content analysis process of interview

After selecting the analysis method and designing the coding list, an analytical process was considered. Magnusson and Marecek (2015) provide step-by-step guidance to find meaning in people's talk, transcribing the interviews, coding and synthesising the shared meanings. I refined the analytical procedures utilising Nvivo, a feasible instrument to systematically highlight and visualise similar contents under the codes and categories.

Step 1: Transcribe recordings into a textual form. I transcribed all the 17 Chinese scripts, sent them to interviewees, and asked for their feedback to check whether there was any ambiguity. After confirming the accuracy, I translated the Chinese written

texts into English and asked two other colleagues fluent in both languages to check whether the translation reflected the interviewees' original meaning.

Step 2: Input all the revised English transcripts in Nvivo.

Step 3: Create a final coding list (Appendix G2) for data analysis according to the initial coding list (Appendix G1) and analysis of students' responses.

Step 4: Code each transcript by recognising the shared information as sub-codes directly or indirectly relevant to the anticipated category and labelling them in different colours.

Step 5: Classify these sub-codes into upper-level codes with the same topics.

Step 6: Reread the interview transcripts, modifying and determining the codes for the second coding cycle. New codes or categories may be added according to interviewees' responses that I did not foresee. All codes and categories were reexamined to make sure that there were sufficient data for comparison.

Step 7: Calculate the frequencies of codes and select corresponding quotes for each code to support the results, then summarise each category.

6.6 Ethical Considerations

Given no official ethical guidance or consensus in China, the ethical considerations were mainly based on the Ethical Guidelines for Educational Research released by the British Educational Research Association (BERA) in 2018. Two strands of responsibilities were considered from the perspectives of participants and researchers. Firstly, this case study was conducted in an international field, mainland China, where English is not the language most Chinese customarily use. Therefore, choosing which language was my first consideration. Although participants have learnt English for at least ten years and gained a comparatively high score in Gaokao, it did not indicate all of them were good at speaking English due to their diverse backgrounds. In this sense, I determined to use their native language, Mandarin, to interview, designed the questionnaires in both Mandarin and English and chose the simplified Chinese version of the CT tests to help them better engage in these activities. This option may avoid the potential ethical dilemmas for the participants and me. Since we were Chinese, it would not cause any anxiety for students to fear failure in the mentioned activities. At the same time, as a researcher in this study, I was not worried about communicating with them in Mandarin. When analysing data, I translated it into English to be available for completing my study. This option of collecting data in

participants' native language, translating the Mandarin scripts into English and then analysing data in English makes my research accessible locally and internationally due to the internal setting where English is not prevalent (the 73rd legislation). In addition, I felt safe and secure conducting my research at that university (the 84th legislation) because I have worked there for nearly 20 years. And I adapted to the cultural setting and teaching environment and gained support from my colleagues. Therefore, I did not feel lonely (the 22nd legislation).

Secondly, I informed students' consent by filling out consent forms (Appendix J) before conducting research (the 8th legislation), including the PBL class, CT tests, questionnaires and interviews. Before starting the study, all participants were informed clearly about their rights, such as they could withdraw at any time if they wanted to without giving any reason (the 31st legislation). When conducting questionnaires and interviews, participants were told about the purposes and what they needed to do (the 9th legislation). More importantly, I emphasised that their comments or feedback had nothing to do with their academic performance to ease their anxiety or worries if providing negative voices or responses (the 47th legislation).

Thirdly, all the collected data were stored safely (the 50th legislation) and analysed confidentially (the 40th legislation). During the data collection stage, I recognised their preliminary ownership and kept confidential about their information (i.e. personal information, test results, and perspectives). After data collection, rather than using portable data storage devices such as laptops or USB sticks, I stored them in the university's Onedrive, a secure university online repository with logins controlled by the research project team. This data management method followed the *Research Ethics Code* of the University of Northampton (Research Ethics Committee, 2022). All the collected data and research results were stored in different data files and folders as password protected and checked regularly. At the data analysis stage, the data were monitored to be kept confidential, secure and anonymous. When analysing data and reporting results, I asked their permission to anonymously use excerpts in their essays and interviews to support quantitative results.

These three aspects of ethical considerations reduced the bias arising from my research desire to obtain satisfactory results. It also minimised the possibility of

participants' unreal responses because of being afraid of affecting their academic performance by providing negative comments. As a result, it eliminated the subjective impacts.

6.7 Summary

This chapter first elaborated on how I conducted PBL teaching interventions by presenting teaching procedures. It was followed by a mixed method of what quantitative and qualitative data were collected, the procedure of the pilot study and the formal data collection procedure. Then, the mixed data analytical methods and processes based on different requirements were introduced. Finally, the ethical issues were considered.

Chapter 7 Reliability and Validity of the Research Design

7.1 Introduction

Any research design should follow a logical set of statements, which could be regarded as criteria to evaluate its quality (Yin, 2018). A consensual concept of "trustworthiness" developed by Lincoln and Guba (1985) was regarded as an indicator to meet the "truth" demand for assessing the quality of a case study (Bassey, 1999; Creswell, 2007; Saldaña, 2011). Lincoln and Guba (1985) further conceptualised that concept into four criteria: credibility, transferability, dependability, and confirmability. Merriam and Tisdell (2015) synthesised several researchers' interpretations of such criteria and put forward their understanding:

•Credibility: seeking the correspondence between research and the authentic world by emphasizing multiple possible accounts of social reality;

•Transferability: seeking the possibility of transferring research findings from one group to another;

•Dependability: demonstrating the degree of justifying the conclusions by keeping the systematical trail;

•Confirmability: showing the self-critical attitudes of the researcher(s) about one's perceptions affect the research.

Moreover, some scholars put forward fine-grained and cost-efficient tactics in different research stages for each criterion, such as triangulation or peer review (Bryman, 2016; Merriam and Tisdell, 2015; Yin, 2018). Based on their suggestions, a checklist was designed to evaluate the reliability and validity of this study (Table 7.1). This chapter discusses the above four criteria of credibility, transferability, dependability and confirmability, following the tactic checklist.

Table 7.1 A tactic checklist for enhancing the trustworthiness of the case study

Criteria	Tactic	Research stages		
		data collection	data analysis	research report
Credibility	• triangulation	√	√	√
	• member reflections	√	√	
	• peer review	√	√	
Transferability	• case selection	√		

	• thick description		√	√
Dependability	• triangulation	√	√	
	• member reflections	√	√	√
	• peer review	√	√	
	• audit trail		√	
Confirmability	• ethics	√	√	√

7.2 Credibility

The credibility was suggested to be achieved by triangulation (Bassey, 1999; Cohen *et al*., 2018; Flick, 2017; Yin, 2018), member checking (Creswell, 2007; Duff, 2006) and peer review (Creswell, 2007; Hancock and Algozzine, 2006). In terms of the triangulation, the credibility manifested in stages of data collection, analysis and result report originating using multiple data sources. As previously discussed, a mixed-method design was employed in this study (Chapter 5, section 5.3.4). Specifically, this study collected both quantitative data (i.e. CT tests and questionnaires) and qualitative data (i.e. students' writing projects and interviews) (Chapter 6, section 6.3). In addition, research results were triangulated by combining qualitative and quantitative results. For example, quantitative questionnaire results investigating students' attitudes towards CT-oriented PBL teaching were testified by detailed interpretations from qualitative interview results.

Regarding the member reflections, the credibility was examined by soliciting participants' feedback on the preliminary or emerging findings (Bassey, 1999; Merriam and Tisdell, 2015, Tracy, 2010). Interviewees in this study were invited to check whether my preliminary transcripts of the interview honestly or clearly expressed their perspectives, feelings or attitudes. If not, they had the right to correct factual errors and share their feedback with me.

In terms of peer review, credibility means "providing an external check of the research process" (Creswell, 2007, p.208). Specific to this study, it manifests in designing questionnaires and validating my English translation of the interview. On the one hand, two colleagues were invited to check the questionnaire design as part of my pilot studies (section 6.4.1.2f). I designed two made-to-measure questionnaires because no well-established attitude questionnaire served my specific research purpose. Piloting them by my colleagues was an initial step in enhancing their credibility. A further step that increased their dependability would be discussed in the

section on internal consistency (section 7.4). On the other hand, two colleagues were invited to check whether my English translation accurately expressed the same meaning as the Chinese transcripts for the data collected from the interview. They were English teachers qualified to check my English translation, and their comments laid a foundation for coding data.

7.3 Transferability

Transferability meant the possibility of transferring the research findings to other research with a similar context. Case selection and thick descriptions were two tactics suggested to achieve it (Bryman, 2016; Creswell, 2007; Flyvbejerg, 2006; Merriam and Tisdell, 2015; Tracy, 2010). The first strategy has been discussed in Chapter 5, section 5.3.3.2. The second tactic in this study included descriptions of the context, participants, and the findings with adequate evidence with quotes from participant interviews and students' writing. By providing sufficient illustrations to contextualise this study, readers will be able to clarify the degree of similarity between this research context and their situations and, therefore, transfer the findings in this study to their research. In this sense, this study was expected to be transferred to other studies focusing on developing thinking using CT-oriented PBL in other subjects or different levels of students.

7.4 Dependability

Traditionally, dependability refers to the extent to which findings can be repeated. However, it is problematic in social sciences (Merriam and Tisdell 2015) since human actions are never static. As a result, researchers cannot replicate a study. In addition, researchers may adopt different interpretations of the same data. In this light, dependability in this case study means whether the results are consistent with the data collected. Qualitative data collection instruments were the suggested strategies: interview and writing projects in the present study.

7.4.1 Qualitative data collection instruments

Regarding the interview, my consideration was my lacking interview experience, the time to conduct the interview and the selection of interviewers. Firstly, as an inexperienced interviewer, my questioning skills may not have been sufficient to

create an interactive environment for all the 17 interviewees, which may have negatively influenced the data quality. The second consideration was the selection of interviewers. Although students were free to be interviewed, it would be better to have one from each group in order to learn about their CT performance in different groups. I expressed my desire, stressed the right of interviewees and clarified that their academic performance would not be affected no matter their decisions. Even though I made such an announcement, I suspected some students might be "forced" to participate out of respect for teachers to follow my desire within the Chinese teacher-authority culture. I assumed the recruiting results were normal: two groups were "missing", and two students from the same group volunteered to join the interview.

7.4.2 Quantitative data collection instruments

Besides the qualitative data collection instruments, quantitative data collection tools of the questionnaire and standardised CT tests should discuss as a mixed research design. Regarding the questionnaires, their internal consistency is the most common tactic to testify the reliability (McMillan and Schumacher, 2010), especially for the made-to-measure questionnaires with a specific research purpose and limited resources (Dörnyei, 2009). There are two simple, effective and popular indicators to crosscheck the degree of internal consistency. One is Cronbach's Alpha (coefficient), and the generally accepted value range is from .70 to .95 (Tavakol, 2018, Trobia, 2011). The other is the Mean Inter-item Correlation (Dörnyei, 2009), and the recommended optimal range is .2 to .4 (Pallant, 2020, Tavakol, 2018).

Results of statistical analysis of two indicators in the CT strategies questionnaire (Part I and Part II) were presented (Table 7.2). The Cronbach's Alpha is .89, indicating excellent internal consistency reliability for this scale. This is supported by the Mean of Inter-item Correlations of .36, suggesting comparatively strong relationships among the items. Part III of the questionnaire was not cross-checked here because it investigated students' attitudes towards the effectiveness of each CT strategy. Instead, Part I and Part II aimed to check students' attitudes towards the effectiveness of improving two dimensions of CT. Therefore, those two indicators were crosschecked for Part I and II (Item 1 to 15). This questionnaire was reliable and valid to answer the second research question.

Table 7.2 Reliability analysis for CT strategies questionnaire

Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q1.	.62	.89
Q2.	.55	.89
Q3.	.49	.89
Q4.	.57	.89
Q5.	.62	.89
Q6.	.69	.88
Q7.	.50	.89
Q8.	.51	.89
Q9.	.52	.89
Q10.	.53	.89
Q11.	.56	.89
Q12.	.62	.89
Q13.	.67	.88
Q14.	.44	.89
Q15.	.65	.88
The Mean of Inter-item Correlations = .36		
Cronbach's Alpha for the 22 items =.89		

The results of the statistical analysis of the PBL questionnaire (Table 7.3) showed that the Cronbach's Alpha was .74 (>.70), indicating good internal consistency reliability for this scale, and the Mean of Inter-item Correlations was .32 (.2 -.4), suggesting comparatively strong relationships among the items. PBL questionnaire was reliable and valid in this study to answer the second research question.

Table 7.3 Reliability analysis for PBL questionnaire

Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q1.	.45	.71
Q2.	.46	.72
Q3.	.58	.67
Q4.	.60	.66
Q5.	.57	.71
Q6.	.47	.74
The Mean of Inter-item Correlations = .32		
Cronbach's Alpha for the 9 items =.74		

Regarding the two CT tests of the CCTST and CCTDI, their reliabilities and validities have been discussed in the previous section (section 6.3.1).

7.5 Confirmability

Comfirmability here means the extent to how I, as a researcher, and the findings were free from the influence of my desire or theoretical inclinations, which was the effectiveness of CT-oriented PBL teaching in improving learners' CT. Although being completely objective is impossible (Bryman, 2016), it ultimately depends on me, who should be open-minded and aware of the study and the unfolding results (Thomas and Magilvy, 2011). This dispositional tendency also pertained to the researchers' ethics, which have been previously discussed (section 6.6).

7.6 Summary

This chapter successively used four criteria of credibility, transferability, dependability and confirmability to assess the reliability and validity of the study. Seven corresponding specific tactics, such as triangulation, member reflections, peer review, case selection, thick description, audit trail and ethics, collaborate to promote the rigorous thinking of the study.

PART III FINDINGS AND DISCUSSION

In the preceding methodology part, I presented the rationale of a single-case study research design, a mixed data collection and analysis research method based on the research questions and discussed the reliability and validity. In this part, the findings and discussion of this study were presented. The first three chapters (i.e. Chapters 8, 9 and 10) report the main findings in accordance with three research questions. The following chapter (i.e. Chapter 11) presents the discussion based on the three major findings by discussing the context for CT development and students' CT learning outcomes.

Chapter 8 Findings: CT Learning Outcomes

8.1 Introduction

This chapter presents the findings in response to the first research question, exploring what development EFL students in a CT-oriented made PBL class. As discussed in Chapter 5, quantitative and qualitative data in this research aimed to achieve cross-validation. Firstly, the descriptive statistics of CT tests (CCTST and CCTDI) before and after PBL are presented. Secondly, the findings of a paired sample t-test in SPSS will be given to check if the scores in CT tests are significantly different. Thirdly, qualitative findings from students' writing and interview through content analysis with NVivo will be illustrated.

8.2 CT Skills Development: Results of the CCTST

This section analysed participants' CT scores in CCTST and investigated whether they achieved apparent enhancement in CT skills after PBL. The descriptive statistics of their pre-and-post-PBL CT scores, including an overall score and three scale scores (i.e. Inference, Analysis and Evaluation) as well as the results of the pair samples t-test are illustrated.

8.2.1 Descriptive statistics: CT scores in thinking skills pre-and post-PBL intervention

Table 8.1 illustrates descriptive statistics of the means (M) and standard deviations (SD) for students' overall and core thinking skills before and after PBL. Information for Understanding CCTST Score Report (34-point versions) (2020b) (Appendix C) provides a rubric of the overall score and qualitative descriptions of the score range accordingly. It is ranked five levels: Superior (24-34) indicated superior CT skills, Strong (19-23) showed strong CT skills, Moderate (13-18) represented moderate CT skills, Weak (8-12) demonstrated weak CT skills, and Not Manifested (0-7) revealed a dearth of CT skills. The mean scores of Overall skill increased from 20.4 before PBL to 22.5 after PBL, between the score range of 19-23, indicating participants made progress in their overall thinking skills at a Strong level. That is to say, students excelled in constantly applying reasoning to make their judgements about what they believed or what to do. Their better performance at the Strong level predicted that these new medical students demonstrated great potential for success in their academic medical studies. Suppose if having been undergoing effective CT training constantly, they may achieve a higher Superior CT level and work better in their future workplaces, especially where highly demanded reasoned decision making and thoughtful problem-solving.

Table 8.1 Average scores with standard deviations for Overall and core thinking skills pre- and post-PBL

	Overall	Analysis	Inference	Evaluation
Pre-PBL	20.4±4.0	4.7±1.3	10.8±2.3	4.8±1.8
Post-PBL	22.5±3.4	5.0+.9	11.6±2.1	5.9±1.9

In addition, mean scores of the three core skills also increase (Table 8.1). The same Score Report provides three Strong, Moderate, and Not Manifested levels for each subscale with different score ranges. In terms of the Analysis skill, a Strong level is 5 or more, a Moderate level ranges from 3 to 4, and a Not Manifested level is from 0 to 2. In terms of the Inference, a Strong level is 12 or more, a Moderate level ranges from 6 to11, and a Not Manifested level ranges from 0 to 5. In terms of the Evaluation, 8 or more is at a Strong level, 4-7 is Moderate, and 0-3 is Not manifested.

The mean score of Inference increases from 10.8 to 11.6. It is the only dimension upgrading to a higher level from a Moderate (*M=10.8, SD=2.3)* before PBL to a Strong level (*M=11.6, SD=2.1*) after PBL. This higher level indicated that they underwent a comparatively huge transformation from having difficulties to finally achieving to draw logical conclusions with more accurate information using diverse reasoning methods.

In comparison, the mean score of Analysis increases from 4.7 to 5.0, and Evaluation increases from 4.8 to 5.9 with the same Moderate level after PBL. The development in Analysis is within the same level, indicating that students improved this skill to some degree but still had difficulty identifying the most relevant information, such as assumptions, reasons or evidence, when making arguments or offering explanations. Similarly, students' Evaluation was also within a Moderate level, indicating that sometimes they might feel difficulty assessing the quality of arguments or the credibility of other people's perspectives, although they improved their Evaluation skills.

8.2.2 The paired samples t-test for CT skills development

Distribution analysis

Distribution analysis was conducted first to examine whether data was distributed normally before conducting paired samples t-test. The normal curves displayed in the following figures revealed that the overall score (Figure 8.1) and subscale scores (Inference, Analysis and Evaluation) (Figure 8.2) were distributed normally before and after PBL intervention.

Figure 8.1 Distributions of Overall CT scores before and after PBL

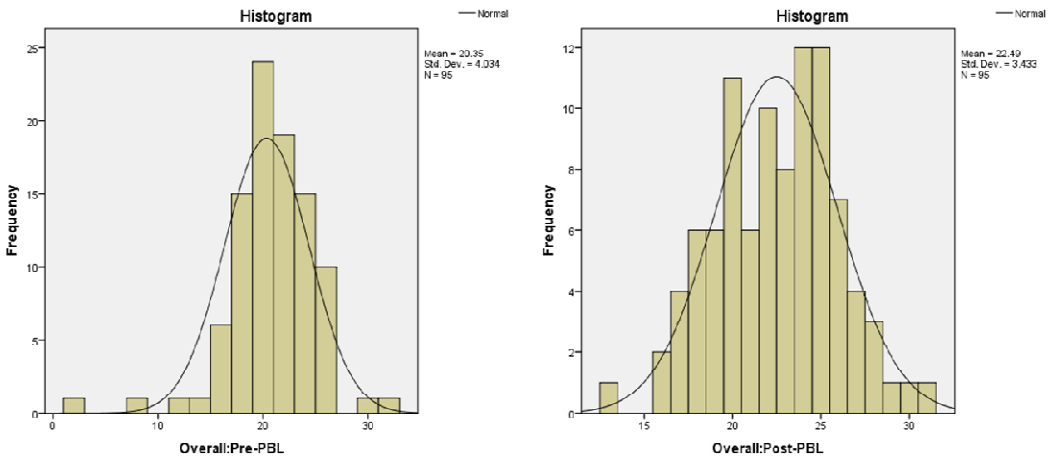
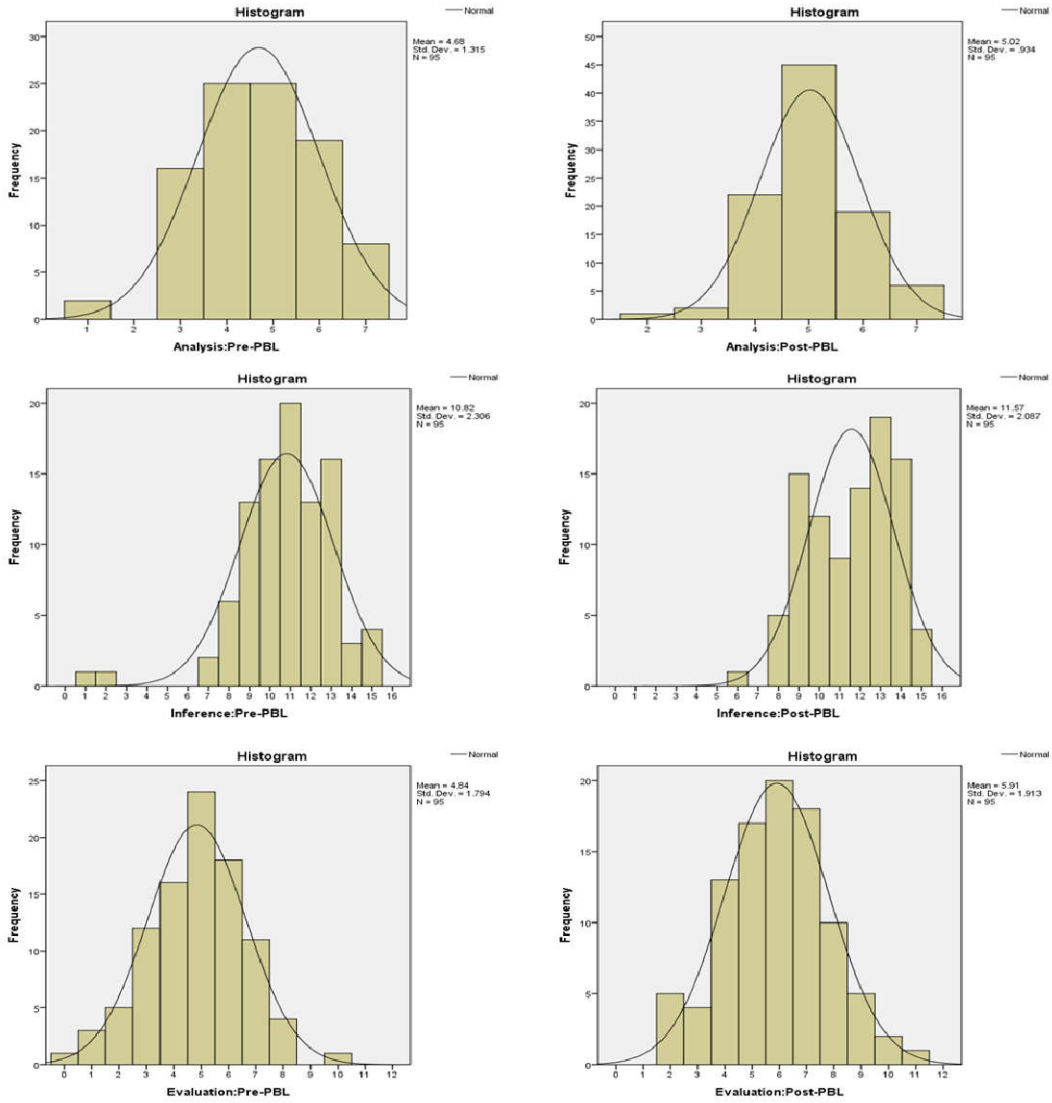


Figure 8.2 Distributions of three core CT skills before and after PBL



Results of the paired samples t-test

Table 8.2 below illustrates results of the paired samples t-test. The mean score of Overall skill increases by 2.15 points after PBL, and the P value (*p=.000*) is less than .05, indicating a significant difference. This P value revealed students' overall critical thinking made a significant development but failed to disclose the degree of the difference (Sullivan and Feinn, 2012). Therefore, the effect size, an independent measurement to quantify the difference between two scores (Creswell and Guetterman, 2018), is necessary to report. The effect size of the Overall score is value of .57, indicating the score after PBL was 0.57 standard deviation higher than the one before

PBL, representing a moderate effect (Table 8.3). In other words, the magnitude of the development was moderate.

Table 8.2 A paired samples t-test of the mean scores of the CCTST

CCTST		M	SD	95% CI	t(df)	P(2-tailed)	d
Overall	Pre-PBL-	-2.15	3.27	-2.81,-1.48	-6.41(94)	.000	.57
	Post-PBL						
Analysis	Pre-PBL-	-.34	1.35	-.61, -.06	-2.43(94)	.017	.30
	Post-PBL						
Inference	Pre-PBL-	-.75	2.21	-1.20, -.30	-3.30(94)	.001	.34
	Post-PBL						
Evaluation	Pre-PBL-	-1.06	1.91	-1.45, -.67	-5.42(94)	.000	.57
	Post-PBL						

Table 8.3 Effect size calculation (Cohen *et al*., 2018, p. 746)

Effect size	
0-0.20	Weak effect
0.21-0.50	Modest effect
0.51-1.00	Moderate effect
>1.00	Strong effect

This overall mean score's significant improvement confirmed previous studies. For example, Eftekhari *et al.* (2016) claim that students significantly improve their Overall CCTST mean score by drawing argument maps in 12 weeks. Another study by Butchart and Forster (2009) prove that 43 undergraduates obviously increase their mean score of the post-test of CCTST by applying argument mapping to automatic feedback in 12 weeks' instruction. However, some studies present diametrical results. D'Antoni *et al.* (2010) report 66 first-year medical students made no apparent difference between pre-and post- mean total scores. One justification is eight weeks are too short for students to create mind maps since CT cultivation takes time to develop. Similarly, Bixler *et al.* (2015) conclude that developing thinking skills in a selective course in 4 weeks is unreasonable. In this sense, students' thinking development and teaching time seem to confirm a direct proportion: the longer a teaching length, the higher the possibility of success in CT development. Niu *et al.* (2013) and Cáceres *et al.* (2020) prove that a teaching intervention of more than 12 weeks is more effective than a shorter one. Therefore, CT instruction length seemed to be a key factor determining whether students made a significant CT development.

However, Abrami *et al*. (2015) further point out that the duration of the teaching intervention does not significantly affect the effect sizes of CT teaching results. It means more than 12 weeks of CT teaching may apparently improve students' critical thinking, but it does not guarantee that the longer teaching interventions, the more significant CT development.

Specifically, the mean score of the Evaluation increases by 1.06 points *(p=.000, Cohen'd=.57)* (Table 8.2), and its P value is less than .05, indicating a significant difference. The effect size is .57, revealing the score increased 0.57 standard deviation after PBL. This change represented a moderate effect size (Table 8.3). Scores on Evaluation increased significantly with a moderate degree. In comparison, the mean scores of the Inference *(p=.001, Cohen's d=.34)* and Analysis *(p=.017, Cohen's d=30)* increase by .75 and .34 points. Both P values are less than .05, indicating significant differences. Their effect sizes are .34 and .30 individually (Table 8.2), revealing the scores increased 0.34 and .30 standard deviation after PBL. Both changes are within the 0.20-0.50, indicating a modest effect size (Table 8.3). Scores on Inference and Analysis improved significantly after PBL with a small magnitude.

In a sentence, the CCTST results revealed that students' Overall and all the three targeted core skills significantly developed after PBL. Furthermore, Overall skill and Evaluation changed more than Inference and Analysis. However, Inference was the only core thinking skill that updated from a Moderate to a Strong level, even though it developed to a modest degree.

8.3 CT Dispositions Development: Results of the CCTDI

This section analysed participants' attitudes towards CT and investigated whether the attitudes significantly differed after PBL classes. First, the descriptive statistics of six dimensions of CT dispositional scores before and after PBL were presented. In addition, the results of pair samples t-test analysis of scale scores were illustrated.

8.3.1 Descriptive statistics: CT scores in thinking dispositions pre-and post-PBL

The following table (Table 8.4) presents the descriptive statistics of the means (M) and standard deviations (SD) for students' CT dispositional scores before and after PBL. Generally speaking, students improve their mean scores in five out of six

dimensions: Open-mindedness *(M=37.7+1.4)*, Inquisitiveness *(M=49.6+1.8)*, Analyticity *(M=40.8+2.3)*, Confidence in Reasoning *(M=42.1+3.7)* and Maturity of Judgement *(M=42.6+0.7)*. However, Truth-seeking slightly decreases *(M=31.0-0.4)*.

Table 8.4 Average scores with standard deviations for six subscales pre-and-post-PBL

	Truth-seeking	Open-Mindedness	Inquisitiveness	Analyticity	Confidence in Reasoning	Maturity of Judgment
Pre-PBL	31.0±5.2	37.7±4.0	49.6±4.6	40.8±4.6	42.1±6.1	42.6±5.0
Post-PBL	30.6±4.9	39.1±3.8	51.4±4.9	43.1±4.9	45.8±6.7	43.3±5.8

Based on the CCTDI numerical score range and qualitative interpretations provided by the CCTDI User Manual (2021) (Appendix D), test-takers' attitudes towards six mindset attributes were divided into five types, including Strong Positive (50-60), Positive (40-49), Inconsistent/Ambivalent (30-39), Negative (20-29) and Strong Negative (10-19). Firstly, participants held an ambivalent attitude towards Open-mindedness (*M=37.7 before PBL* and *M=39.1 after PBL*) and Truth-seeking (*M=31.0 before PBL* and *M=30.6 after PBL*) before and after PBL. In particular, the mean score for Truth-seeking was the lowest, and it was the only one whose score reduced slightly after PBL among all the subscales. This result is consistent with the research of Bixler *et al.* (2015), reporting the lowest score of Truth-seeking for 27 fourth year medical students. The unchanged ambivalent attitude of the two dimensions suggested that students were still uncertain whether they could open their minds or break through stereotypes when acquiring new knowledge or making decisions after PBL.

Secondly, they maintained a positive attitude towards Analyticity (*M=40.8 before PBL* and *M=43.1 after PBL*), Confidence in reasoning (*M=42.1 before PBL* and *M=45.8 after PBL*) and Maturity in judgement (*M=42.6 before PBL* and *M=43.3 after PBL*) before and after PBL, revealing that they kept confident in analysing the information and applying reasoning skills to make reasonable judgements after PBL.

Finally, they held the same strong positive attitude towards Inquisitiveness (*M=49.6 before PBL* and *M=51.2 after PBL*) before and after PBL, revealing that students kept their strong curiosity about acquiring knowledge. At the same time, they believed they employed sufficient capacities to explore the world before and after PBL.

8.3.2 Inferential statistics: results of the pair samples t-test

Distribution analysis

Distribution analysis was conducted first to examine whether data was distributed normally before conducting paired samples t-test. The Kolmogorov-Smirnov test produced test statistics that were used for normality. P values (quoted under Sig. for Kolmogorov-Smirnov and Shapiro-Wilk) of Six subscales of CT disposition are more than .05 (Table 8.5, see below) before and after PBL. These non-significant results (p >.05) indicated normal distributions (Pallant, 2016). Therefore, paired samples t-tests were employed to assess whether there were significant differences in sub-scales of CT disposition after PBL.

Table 8.5 Tests of normality of six subscales of CT disposition before and after PBL

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
Truth-seeking:	Pre-PBL	.096	93	.036	.984	93	.310
	Post-PBL	.092	93	.050	.980	93	.162
Open-	Pre-PBL	.100	93	.022	.971	93	.035
mindedness:	Post-PBL	.110	93	.007	.974	93	.055
Inquisitiveness:	Pre-PBL	.087	93	.080	.982	93	.243
	Post-PBL	.097	93	.030	.966	93	.017
Analyticity:	Pre-PBL	.096	93	.035	.980	93	.162
	Post-PBL	.113	93	.005	.976	93	.078
Confidence in	Pre-PBL	.090	93	.062	.980	93	.160
Reasoning:	Post-PBL	.079	93	.199	.984	93	.310
Maturity of	Pre-PBL	.102	93	.018	.980	93	.154
Judgment:	Post-PBL	.070	93	.200*	.954	93	.002

*. This is a lower bound of the true significance. a. Lilliefors Significance Correction

Results of the paired samples t-test

The paired samples t-test showed that students significantly improved four out of six mindset attributes (Table 8.6). Firstly, the mean score of Open-mindedness increases by 1.45 points (*p=.000, Cohen's d=.37*), Inquisitiveness increases by 1.85 (*p=.000, Cohen's d=.39*) and Analyticity increases by 2.30 (*p=.000, Cohen's d=.48*) after PBL. Their P values are less than .05, indicating significant differences. In addition, their effect sizes are 0.37, 0.39 and 0.48, showing the mean scores after PBL were 0.37, 0.39 and 0.48 standard deviations higher than the ones before PBL and their effect

sizes are modest (Table 8.3). Secondly, the mean score of Confidence in Reasoning increases by 3.65 points (*p=.000, Cohen's d=.57*) and the P value is less than .05, indicating a significant difference. Furthermore, the effect size is 0.57, revealing the score after PBL was 0.57 standard deviation higher than before PBL with a moderate effect size (Table 8.3).

In a sentence, students' attitudes significantly changed the subscales mentioned above after PBL. Among the four mindset attributes, their attitudes towards Confidence in Reasoning changed more remarkably than the other three attributes of Open-mindedness, Inquisitiveness, and Analyticity.

Table 8.6 Paired samples t-test of the CCTDI Score before and after PBL

CCTDI		M	SD	95% CI	t(df)	P(2-tailed)	D
Truth-seeking	Pre-PBL-	.31	5.14	-.75, 1.37	.59 (92)	.560	.06
	Post-PBL						
Open-mindedness	Pre-PBL-	-1.45	4.65	-2.41, -.49	-3.01(92)	.003	.37
	Post-PBL						
Inquisitiveness	Pre-PBL-	-1.85	4.42	-2.76, -.94	-4.04(92)	.000	.39
	Post-PBL						
Analyticity	Pre-PBL-	-2.30	4.50	-3.22,-1.36	-4.91(92)	.000	.48
	Post-PBL						
Confidence in Reasoning	Pre-PBL-	-3.65	5.27	-4.73,-2.56	-6.67(92)	.000	.57
	Post-PBL						
Maturity in Judgement	Pre-PBL-	-.61	6.07	-1.86, .64	-.97 (92)	.333	.11
	Post-PBL						

In contrast, no significant improvements were found in another two mindsets of Truth-seeking (*p=.56, Cohen's d=.06*) and Maturity in judgement (*p*=.33, *Cohen's d=.06*) (Table 8.6) after PBL. Both P values are more than .05, indicating insignificant differences. Meanwhile, their effect sizes are .06 and .11, showing weak effects sizes (Table 8.3). This change revealed that the magnitude of differences in the mean scores was too small to be neglected.

In brief, the CCTDI revealed that students significantly enhanced in four dimensions of CT disposition: Open-mindedness, Inquisitiveness, Analyticity and Confidence in Reasoning and no significant improvements in Truth-seeking and Maturity in Judgement after PBL. Considering an insignificant improvement and the ambivalent

attitude with the lowest mean score students held, Truth-seeking was the weakest scale. In contrast, Inquisitiveness was the strongest with significant development and the highest mean score. In addition, Open-mindedness is the second weakest scale because students' attitude is still ambivalent, although they made statistical development. Finally, students made the most significant improvement in Confidence in reasoning with the largest effect size.

8.4 The Content-specific CT Skills Development: Results of Students' Writing Analysis

This section explores whether frequencies of thinking skills in writing increase significantly in students' final drafts. Firstly, descriptive statistics will be presented, including three categories with ten codes in students' first and final drafts. Secondly, the results of Wilcoxon's signed rank tests will be presented with specific examples from students' written texts. As discussed in chapter 6 (section 6.5.3.1), the content analysis with Nvivo was based on a rubric for assessing CT skills in writing, including three targeted categories (i.e. Inference, Analysis and Evaluation skills) and corresponding ten items with acceptable and weak scales (Appendix I). The texts scaled as acceptable will be coded according to ten items (Table 8.7).

Table 8.7 Categories and codes of students' writing

Categories	Codes
1. Inference	code 1: using cohesive devices
	code 2: drawing logical conclusions
	code 3: using multiple sources
2. Analysis	code 4: writing main claims
	code 5: writing sub-claims
	code 6: writing thesis statements
	code 7: using comparison and contrast
3. Evaluation	code 8: determining what n;eeds doing
	code 9: providing implications
	code 10: evaluating whether claims and evidence were convincing or valuable

8.4.1 Descriptive statistics: results of frequencies of using CT skills in writing

In total, 56470 words in 110 essays were analysed, including 55 first drafts (S01-1-S55-1) and final drafts (S01-2-S55-2). S01 means the No. 01 student; -1 represents the first draft, and -2 represents the final draft. Therefore, S01-1 represents

the first draft of the No.01 student, while S01-2 represents the final draft of the No.01 student. Figure 8.3 below reveals students increased the frequencies of 9 out of 10 codes in their writing after PBL except for the 7th code (i.e. comparison and contrast). This growing trend indicated that students increased their ability and consciousness to use most CT skills, enhancing their thinking performance in writing.

Figure 8.3 Frequencies of CT skills in writing before and after PBL

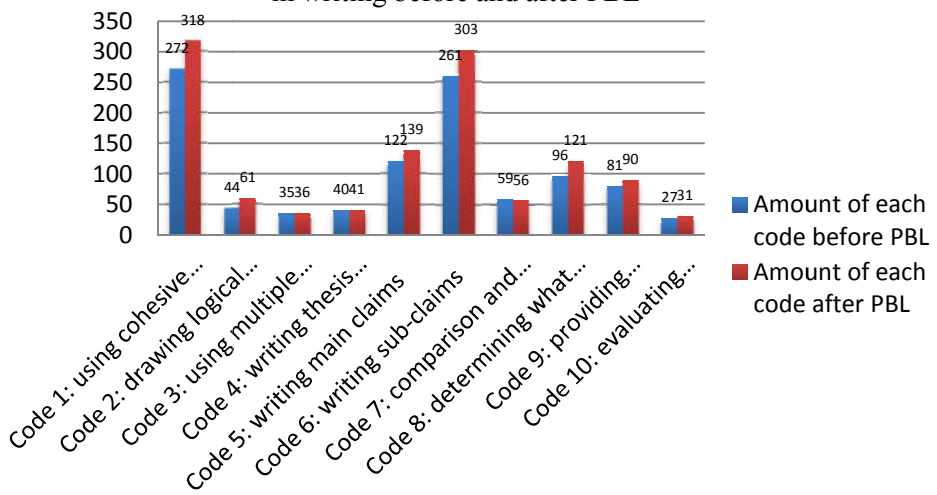
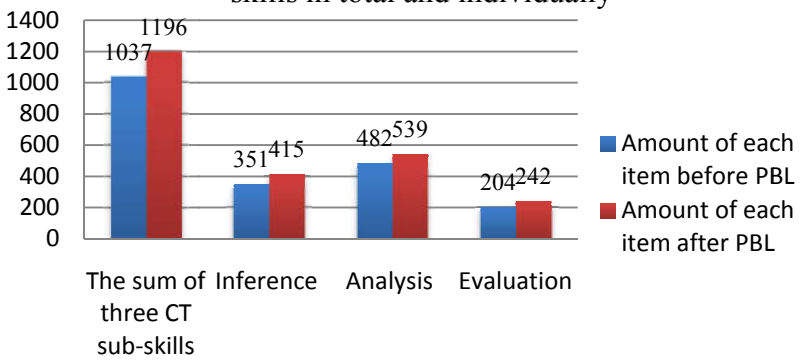


Figure 8.4 below reveals the same growing trend for the sum frequencies of the total three categories and each category of Inference skill (i.e. the sum of code 1-3), Analysis skill (i.e. the sum of code 4-7) and Evaluation skill (i.e. the sum of code 8-10). Among the three core skills, the Evaluation skill has the highest growth rate (15.7%), followed by Inference (15.4%) and the last one is Analysis (10.6%).

Figure 8.4 Frequencies of core thinking skills in total and individually



8.4.2 Inferential statisitcs: results of the Wilcoxon signed rank test

Because thinking writing skills frequencies were less likely distributed normally, a Wilcoxon signed rank test, rather than a paired samples t-test, was carried out to assess whether frequencies significantly increased after PBL. If they did, it meant students would develop their CT skills in writing.

Table 8.8 below reports that the P values of the Inference (*p=.000, Cohen's d=.36*), Analysis (*p=.003, Cohen's d=.29*) and Evaluation (*p=.001, Cohen's d=.32*) are less than .05. Therefore the frequencies of such writing skills made significant differences after PBL. In addition, their effect sizes range from 0.21 to 0.50, indicating that the differences were modest (Table 8.3). Their less than .05 P values and small figures of effect sizes suggested that students made significant development in all three targeted skills, but their development magnitudes were small. This result is consistent with the apparent development of all the three CT skills in the CCTST of the current study.

Table 8.8 A Wilcoxon signed rank test of the frequencies of three core thinking skills in writing during and after PBL

	Inference: post-PBL- pre-PBL	Analysis: post-PBL- pre-PBL	Evaluation: post-PBL-pre-PBL
Z	-3.817[b]	-3.004[b]	-3.347[b]
Asymp. Sig.(2-tailed)	.000	.003	.001
D	.36	.29	.32

a. Wilcoxon Signed Ranks Test b. Based on negative ranks.

8.4.2.1 Inference skill development

Three items were coded in Inference: cohesive devices (Code 1), drawing logical conclusions (Code 2), and using multiple sources (Code 3) (Table 8.7). The Wilcoxon signed rank test result indicates that their P values are below .05 (Table 8.9), indicating the corresponding frequencies made significant differences. In addition, their effect sizes are in the range of 0.20-0.50, representing a modest effect size. Therefore, students made significant development in all three codes of Inference. However, the difference was small.

Table 8.9 The Wilcoxon signed rank test of frequencies of three codes of Inference in writing during and after PBL

	Cohesive devices: post-PBL -pre-PBL	Logical conclusions: post-PBL - pre-PBL	Multiple sources: post-PBL - pre-PBL
Z	-3.130[b]	-2.676[b]	-2.496[b]
Asymp. Sig.(2-tailed)	.002	.007	.013
D	.30	.25	.25

b. Based on negative ranks.

The significant enhancement of using cohesive devices (code 1)

The significantly increased frequencies of code 1, using cohesive devices in writing, indicated that students strengthened their abilities to use more accurate linking words or phrases to display logical or coherent relationships in writing after PBL. Their CT development in writing was summarised into two types through analysing and comparing students' first and final drafts related to this code. Firstly, students used more cohesive devices to meet the needs of increased statements when justifying their points of view. For example, they provided more claims to elaborate thesis statements or sub-claims to support the main claims (S01). Or they offered additional comments to explain further or evaluate the previous statements (S04).

> Example 1: College dating: yes or not? (Unit 4)
>
> **The final draft** (S01-2): In that case, how can we set up a sensible viewpoint in the university? (The main claim). <u>First of all</u>, we are supposed to have a common belief and persist and then improve ourselves (1st sub-claim). <u>Besides</u>, I think we should handle this relationship between dating and study or work correctly (2nd sub-claim). <u>Moreover</u>, we must know that college dating may not always be delighted (3rd sub-claim). <u>Last but not least</u>, ... so, we also should increase our communication skills and expand our circle of friends (4th sub-claim).

S01 added one more main claim and four new corresponding claims in the final draft. Accordingly, he/she used four more transitional signs to connect four parallel sub-claims of suggestions to support that new main claim.

> Example 2: GaoKao: Never Give Up (Unit 2)
>
> **The first draft** (S04-1): ... Last but not least, insistence is indispensable. Only with continual insistence, it's more possible to be admitted to a better university."
>
> **The final draft** (04-2): ... Last but not least, insistence is

indispensable. Only with continual insistence, it's more possible for us students to be admitted to a better university. But you need to make a balance between study and recreation. That is significant because it's harmful to our body and mind to study for a long time without a short break.

S04, in the final draft, provided additional remarks by further reminding what students should pay attention to after providing three sub-claims of attitudes towards taking the college entrance examination. The adversative conjunction "But" emphasised the necessity of balancing between hard-working and the rest. In addition, this student further evaluated the importance of that necessity by providing a reason. In this case, "because" as a conjunction of expressing cause was used.

Secondly, students made clearer logical minds by using more cohesive devices to link sentences or paragraphs to progress from one idea to another. Using signal words helps writers clarify how various claims are connected in an argument or the reasoning process (Cottrell, 2017). In addition, it was more convenient for EFL students to use such explicit connectives to help them present the reasoning process in English writing. As a result, they built up more coherent relationships such as progression, transition, causality or compare and contrast. The following examples suggest two ways: adding new linking words or correcting the original misused cohesive devices (S06 and S33).

Example 3: Differences between the Old and the Young (Unit 3)

The first draft (S06-1): The old generation grew in poverty, war and disaster. The young generation grew in a time of peace and affluence. This leads to the two generations having different ideas.

The final draft (S06-2): The old generation grew in poverty, war and disaster while the young generation grew in a time of peace and affluence. This leads to the generations having different ideas.

S06 used the conjunction "while" to contrast the difference between the young and old generations in the final draft. It was obviously more coherent than the first draft without such a cohesive device.

Example 4: The management of money (Unit 1)

The first draft (S33-1): Most parents pay too much attention

to their children's achievements, but ignore the application in daily life, and there is no place where English can be used in life to exercise their English ability so that English out of life lost the original meaning of learning English.

The final draft (S33-2): Most parents pay too much attention to their children's achievements, but ignore the application in daily life, and there is no place where English can be used in life to exercise their children's ability. As a result, English is out of life, losing the original meaning of learning English.

The above example shows that S33 replaced "so that" with "as a result" in the final draft to establish a causal relationship between the problem of English learning and being far away from the original purpose of English learning. It seems that S33 has developed his/her logical thinking skills.

The significant development in drawing logical conclusions (code 2)

The significantly increased frequency of code 2, drawing logical conclusions, revealed that students tended to think more logically by less making simple summaries. Firstly, students raised their awareness of making logical conclusions based on evidence or gathering information (S07).

Example 5: The key to filling the generation gap (Unit 3)

The first draft (S07-1): ... Due to these inappropriate behaviours, we, who are in adolescence, not only make the generation gap more and deeper but also hurt our parents' hearts.

The final draft (S07-2): ... Due to these inappropriate behaviours, we, who are in adolescence, not only make the generation gap more and deeper but also hurt our parents' hearts. Children ought to understand their parents' good intentions. (logical conclusion)

S07 simply summarised some inappropriate behaviours into a shorter overall summary without any judgement or conclusion of a deep generation gap in the first draft. In contrast, he/she provided a logical conclusion that "children ought to understand their parents' good intentions" after summarising those behaviours in the final draft.

Secondly, the conclusions were more related to the thesis statement. Students may realise the necessity of making a logical conclusion but sometimes neglect a

fundamental principle of being closely related to the thesis statement. Some final drafts showed students' enhancing ability or consciousness to solve this problem (S16).

> Example 6: Be active in English Class (Unit 1)
>
> **The first draft** (S16-1): " To conclude, fifty-five minutes in class is limited. What we must be aware of is that it belongs to both teachers and students. Sitting there and accepting knowledge passively is not the final aim of teaching, which is the least efficient way to learn. Only when we listen, think, ask, communicate ideas and digest actively, can we have a better realization of what we learn.
>
> **The final draft** (S16-2): " To conclude, forty-five minutes in English class is limited. What we must be aware of is that it belongs to both teachers and students. It is us that is the leading role of class. More significantly, the majority of the English language is to speak, namely, be used to communicate. Only do we think, ask, speak, communicate ideas and challenge ourselves bravely and actively, can we have a good command of English.

S16's conclusion in the final draft was more directly linked to the thesis of the English class than the first draft, which seemed focused on a general subject instead of the English subject. At the same time, this student explicitly pointed out that they should act a leading role in the classroom and be active in class

The significant development of using multiple sources (code 3)

The apparent improvement in the amount of code 3, using various sources in writing, indicated that students tended to use more evidence to support claims from multiple sources (S24, see below). S24 used statistics to support the argument for a declining number of history majors. This example displayed his/her willingness to use diverse evidence to interpret claims after PBL.

> Example 7: Do you know humanities? (Unit 2)
>
> **The final draft** (S24-2): Now(a)days, as more and more people enter university for education, the proportion of students willing to devote themselves to the humanities is declining. A study of Harvard students from 2008 to 2016 found that the number of history majors dropped from 231 to 136.

8.4.2.2 Analysis skill development

Four items were coded in Analysis: providing main claims (code 4), offering sub-claims (code 5), writing thesis statements (code 6) and making comparisons and contrast (code 7) (Table 8.7). Results of the Wilcoxon signed rank test present that the P values of code 4, providing main claims (*p=.006, Cohen's d=.26*) and code 5, providing sub-claims (*p=.001, Cohen's d=.30*), are less than .05 (Table 8.10), indicating both frequencies made significant differences after PBL. In addition, both effect sizes are in the range of 0.20-0.50, showing a modest effect size (Table 8.3). Besides, the P values of the other two codes are more than .05 (Table 8.10), including code 6, writing thesis statements (*p=.796, Cohen's d=.02*) and code 7, comparison and contrast (*p=.914, Cohen's d=.08*), indicating insignificant differences after PBL. In addition, both effect sizes are in the range of 0-0.20, showing a weak effect (Table 8.3). The difference magnitudes were too small to be noticed.

Table 8.10 The Wilcoxon signed rank test of the frequencies of four items of Analysis in writing during and after PBL

	Thesis statements: post-PBL-pre-PBL	Main claims: post-PBL-pre-PBL	Sub-claims: post-PBL-pre-PBL	Compare and contrast: post-PBL-pre-PBL
Z	-.258[b]	-2.736[b]	-3.197[b]	-.108[b]
Asymp. Sig. (2-tailed)	.796	.006	.001	.914
D	.02	.26	.30	.08

b. Based on positive ranks.

The significant development of providing main claims (code 4)

The significantly developed frequencies of code 4, providing main claims in writing, indicated students put forward more accurate central claims to support the thesis statements than in the first draft. Firstly, students increased the frequencies of statements to strengthen the thesis from more angles, following a logical thinking process (S02).

Example 8: Contemporary Consumer (Unit 5)

The first draft (S02-1): ...<u>The consumption concept of</u> people who live at this age has a huge change. (1st claim) ... <u>So what we should do to escape this trap?</u> (2nd claim)

The final draft (S02-2): ...<u>The consumption concept of</u> people who live at this age has a huge change. (1st claim)

> What is the reason for this? (2nd claim) Today, as this phenomenon expands in China, I think it should be given great importance. (3rd claim) What we should do to escape this trap? (4th claim)

S02 in the final draft supplemented two more main claims: the reasons and the drawbacks of the consumption concept change. Therefore, this essay became more logical by describing a problem, analysing why it arose, identifying its negative influence and making suggestions to avoid it. While in the first draft, he/she expressed the perspectives from description straightly to advice without any reasoning process. This thinking gap-filling process indicated that S02 became more analytical in writing.

Secondly, some students added a topic sentence as the main claim to summarise the sub-claims at the beginning of a paragraph (S43). S43 in the final draft started that paragraph by adding a topic sentence. Following this sentence, the readers could easily catch up with the main message: why dumb English was common in China. While in the first draft, this student directly provided two reasons without any topic sentence, and this deficiency may confuse the readers in figuring out the main point. In this sense, S43 increased his/her intention to write a topic sentence as a signpost to start a new paragraph.

> Example 9: Dumb English (Unit 1) (dumb English)
>
> **The first draft** (S43-1): First of all, in China, parents or teachers make too much account of students' grades. ... Secondly, the skills of listening and speaking are ignored completely.
>
> **The final draft** (S43-2): The reason for the phenomenon of dumb English is not unique. First of all, in China, parents or teachers make too much account of students' grades. ... Secondly, the skills of listening and speaking are ignored completely.

(Dumb English means a dilemma that students cannot speak English but are good at reading, listening or writing.)

The significant development of using sub-claims (code 5)

The obviously increased frequencies of code 5, offering more sub-claims, indicated students put forward more related sub-claims to support main claims and analysed

their points of view in writing. Firstly, students wrote more sub-claims because of increased new main claims (S16).

Example 10: Be active in English class (Unit 1)

The first draft (S16-1): Someone might have the question why Chinese students are so fearful of making mistakes. (the main claim) On the one hand, Chinese students undertake enormous pressure from teachers and parents who set high goals. (sub-claim 1) On the other hand, the whole education and examination system in China makes the atmosphere in schools very depressive, hidebound and stiff. (sub-claim 2)

The final draft (S16-2): Someone might have the question why Chinese students are less active in English classes. (the main claim) For one thing, plenty of Chinese students lack confidence. (sub-claim 1) For another, the traditional English teaching model in China pays too much attention on teachers and ignores the leading position of students. (sub-claim 2) At the same time, the dull atmosphere in the English class makes teachers and students lack the necessary communication in the learning process. (sub-claim 3)

S16 in the first draft deviated from the thesis of being active in English class to the thesis of why students were afraid of making mistakes. Therefore, that claim was not coded. Noticing this error, this student writer altered the original main claim to analyse why students were not active in class, and this new idea was directly consistent with the thesis. This change indicated that he/she evaluated the pertinence between the thesis statement and the main claim and then made a judgement to replace the original one.

In the second place, students provided more sub-claims to support the original central claims from more angles (S04). S04, in the final draft, added the third reason to explain why students felt exhausted when facing Gaokao from a new perspective of teachers and parents. This new perspective revealed that this student extended their thinking in the final draft.

Example 11: GaoKao: Never Give Up (Unit 2)

The first draft (S04-1): That is the stress they face. (1st sub-claim) On the other hand, ... this examination is full of selection. (2nd sub-claim)

The final draft (S04-2): "That is the stress they face. (1st sub-claim) On the other hand, ... this examination is full of selection (2nd sub-claim) Teachers' and parents' expectations give us a lot of pressure, which is the other main source of stress. (3rd sub-claim)

In the third place, students provided more precise sub-claims in the final draft (S11). In this example, when analysing why students did not like to learn grammar, S11 stated vaguely by saying "a little difficult" in the first draft, confusing the reader about what exactly challenged them. In the final draft, this student elaborated explicitly on the obstacle of learning grammar: memorising the grammar rules. Therefore, S11 increased the frequency of sub-claims by clarifying the statements and reducing the vagueness.

> Example 12: Learning English grammar (Unit 1)
>
> **The first draft** (S11-1): But nowadays, the situation is not rosy. ... Moreover, it is a little difficult for them to learn grammar well, so they just concentrate on vocabulary and ignore the grammar.
>
> **The final draft** (S11-2): But nowadays, the circumstance is not rosy. ... Moreover, it is not so easy for them to recite grammar rules, so they concentrate on vocabulary instead.

In short, the improvement of main claims and sub-claims was pertinent to the higher quality and growing quantity of statements. This result revealed that students were more aware of the connection between the thesis statements and main claims or between the main claims and sub-claims, especially when intending to add new main claims or substitute original ones.

The insignificant development of code 6: providing thesis statements

Students made no apparent improvement in the frequency of code 6, delivering thesis statements, revealing that they did not significantly develop their ability to state a topic clearly. The final drafts (n=12) that failed to provide valid thesis statements were analysed to explore problems when students elaborated thesis statements, given such limited development and the necessity and significance of writing a thesis statement. Nine (out of 12, 75%) students write thesis statements, but they are too general or vague for readers to predict the direction or grasp the main points (S30).

> Example 13: Attitudes to romance on campus (Unit 4)
>
> **The final draft** (S30-2)
> Thesis statement: Nowadays, it's very common to see that some university students fall in love at the campus. Most of the time, they do everything together. And on campus, we all love our male or female friends to reveal to the public.

1st main claim: Some teachers and schools maintain that love will negatively influence university students.
2nd main claim: However, most parents stay neutral on this issue.
3rd main claim: Some students are against campus love.
4th main claim: However, I hold an opposite view.

S30 had a clear mind about what topic he/she should write about according to the title (i.e. the attitudes to romance on campus). However, this student introduced the popularity of love affairs on campus instead of presenting different attitudes towards campus love in general as a thesis statement. It was possible to describe the popularity of campus love. However, it was less relevant to the intended thesis statement. Therefore, this first paragraph failed to point out the writing direction, even though the student elaborated on attitudes in the following four paragraphs. In this sense, this student could not provide a valid thesis statement in the final draft on account of irrelevance to the title.

Another two (out of 12, 17%) students start the essays by directly writing the first main claim, which may result in a deviation from a logical and organized direction to the following writing (S09).

Example 14 The advantages of college students' love (Unit 3)
The final draft (S09-2):
The 1 st paragraph: What is love?
The 2nd paragraph: In today's university campus, the phenomenon of love has been common. ... However, the significance of college love can't be ignored.
The 3rd paragraph: First of all, it can teach us how to love.
The 4th paragraph: "More importantly, we can leave some unforgettable memories in the university.

S09 first defined a love concept, then described the popularity of love affairs on campus and finally analysed three benefits of love for students in the following three paragraphs in the final draft. It seemed that S09 had a clear writing direction in mind, but the student did not write a thesis statement as a signpost at the beginning. If the student had, readers may easily grasp the key view. What S09 needed was to explicitly write a thesis statement as a direction or summary of the main ideas.

Finally, one student misunderstood the topic's keyword, leading to a failure to write the thesis. In brief, almost all these twelve students had an intention to write thesis

statements, although they failed to do so. Considering a comparatively high percentage (75%) of students who did not write clear thesis statements, as mentioned before, their primary challenge was generating a valid thesis statement by pointing out a clear writing direction or a concise summary in their compositions.

The insignificant development of code 7: using comparison and contrast

Students did not apparently increase the frequency of code 7, using comparison and contrast, revealing that they did not make more similarities or differences in their compositions. This unchanging number did not necessarily mean students didn't make improvements. Unlike code 6 of using thesis statements, which was the primary element for writing, code 7 was not essential for every writing topic or genre. Therefore, it was understandable that they did not use more comparison and contrast in their writing.

Generally speaking, although students only increase the frequencies of 2 out of 4 codes, the sum increases obviously. In this light, students developed their Analysis significantly in writing. Given the higher percentage of students who failed to write the thesis statement and its importance, students needed to make efforts to provide valid ones.

8.4.2.3 Evaluation skill development

Three items were coded in Evaluation: determining what needs doing in a given situation (code 8), discussing the implications (code 9), and assessing whether claims or evidence were workable, valuable or convincing (code 10) (Table 8.7). Results of the Wilcoxon signed rank test show that P values of code 8 (*p=.002, Cohen's d=.29*) and code 9 (*p=.039, Cohen's d=.20*) are less than .05 (Table 8.11), indicating the frequency of such codes made significant differences after PBL. Both effect sizes range between 0.20-0.50, showing a modest effect size (Table 8.3). The results revealed that students made obvious development in those two writing skills; however, the development magnitude was small. Besides, the P value of code 10 (*p=.157, Cohen's d=.03*) is higher than .05 (Table 8.11), indicating its frequency did not significantly differ after PBL. The effect size is .03, presenting a weak effect size (Table 8.3). The result of its P value and the effect size revealed that students did not

make any significant difference, and the development degree was too weak to be noticed.

Table 8.11 The Wilcoxon signed rank test of the frequencies of three aspects of Evaluation in writing during and after PBL

	Determine what needs doing: post-PBL -pre-PBL	Provide implications: post-PBL - pre-PBL	Assess whether the claims or the evidence is workable, valuable or convincing: post-PBL- pre-PBL
Z	-3.096b	-2.066b	-1.414 b
Asymp. Sig. (2-tailed)	.002	.039	.157
D	.29	.20	.03

b. Based on negative ranks

The significant development of code 8: determining what needs doing in a given situation

The significantly increased frequency of code 8, determining what needs doing in writing, revealed that students offered more actionable recommendations for solving problems or making progress. Firstly, students tended to provide more related suggestions or recommendations from different angles (S38).

> Example 15: The wealth gap in education can be bridged (Unit 2)
>
> **The first draft** (S38-1): We must take steps to relieve this issue. For example, we can try to balance educational resources and give the low-income family more convenience and more opportunities to get up close. (1st suggestion) ... I think we should start by solving poverty and then give the poor students blaze a trail. (2nd suggestion)
>
> **The final draft** (S38-2): I think we should merit more attention to how to narrow the wealth gap of education. We must take steps to relieve this issue. On an individual level, we should reduce the gap between rich and poor. ... And we should work hard ... and have enough abilities to help poor students complete their academics. (1st suggestion). At the nation and the social level, we can try to balance educational resources and give the low-income family more convenience and more opportunities to get up close. (2nd suggestion) ... I think we should start by solving poverty and then give the poor students blaze a trail (3rd suggestion) that is easier for their life.

S38 wrote three tips to bridge the wealth gap in education from two dimensions: the individual and national level in the final draft. While he/she provided two suggestions focused on one dimension of the national level in the first draft. This change indicated

that S38 extended the perspective from one to two strands, making this draft more comprehensive than the first one.

In the second place, students provided new main claims about suggestions of what needs doing. S40 stated several ways of promoting education from three dimensions: the school, the parents and the students in the first and final draft. Additionally, that student added two more suggestions on the parents' aspect: the parents should spend more time with children and not sign their kids up after school (i.e. the 3rd and 5th suggestion). The other new suggestion was about students' aspect: they should learn for their entire lives (i.e. the 6th suggestion). These added suggestions indicated that S40 reflected on the previous learning experience, applied it to their reflections in the current situation, and made more reasonable judgements.

> Example 16: Let Education Come Alive Again (Unit 2)
>
> **The first draft** (S40-1): As a result, to revitalize education, we should take action to do more. For some schools, they should allow students to have more after-school activities to enrich their spare time. (1st suggestions) What's more, teachers' lectures should be more lively to encourage students to participate. (2nd suggestion) For some families, parents should not give their children too much pressure to study. (3rd suggestion) For ourselves? We should attach more importance to education and pay more attention to putting what they have learned into practice. (4th suggestion)
>
> **The final draft** (S40-2): ...Today, something must be done to improve our education. For schools, they should allow students to have more after-school activities to enrich their spare time and broaden their horizons. (1st suggestion) What's more, teachers' lectures should be more lively to encourage students' participation. (2nd suggestion) For families, parents should let children reasonably study and give more accompany. (3rd suggestion) Then, parents should not give their children much pressure in the study (4th suggestion), and they are not supposed to sign up kids for too many cram schools. (5th suggestions) For students, they should attach more importance to lifelong learning. (6th suggestion) Besides, they should not read too much, and they should pay more attention to putting what they have learned into practice. (7th suggestion)

The significant development of code 9: discussing implications

Students developed significantly in the frequency of code 9, discussing implications, indicating that students made development in discussing the possible effects or results of claims in their writing. This enhancement primarily focused on their increasing

intentions to write possible influences or outcomes of the claims or evidence (S44, see below).

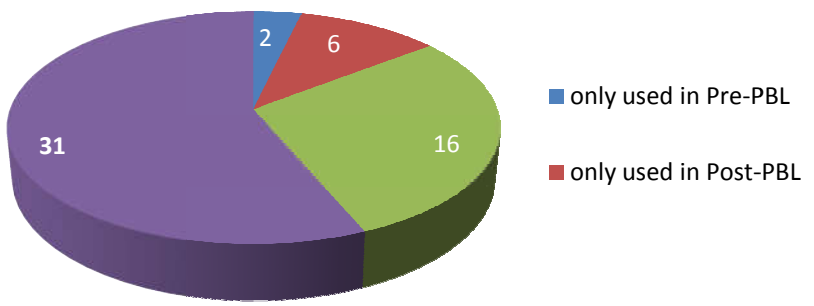
> Example 17: Let Education Come Alive Again (Unit 5)
>
> **The first draft** (S44-1): ...These create a powerful desire and urge people to go shopping. (1st result) .... This thought imperceptibly influenced their children. (1st influence) Their consumption concept has attracted more and more attention, such as running up a credit card and campus's networking platforms loans.
>
> **The final draft** (S44-2): ...These create a powerful desire and urge people to go shopping. (1st result) .... This thought imperceptibly influenced their children. (1st influence) ... Their consumption concept has attracted more and more attention, such as running up a credit card and campus's networking platforms loans, which make students fall into the trap and hurt students deeply. (2nd result)

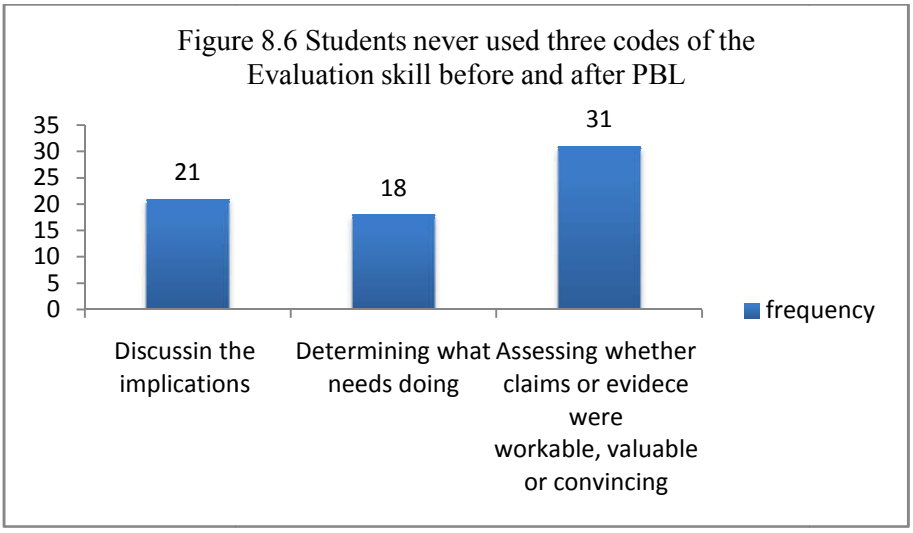
S44 offered one more negative consequence of excessive modification, indicating that he/she thought more profoundly and increased an intention to evaluate the perspectives in writing by exploring implications.

The insignificant development of code 10: assessing whether claims or evidence were workable, valuable or convincing

Students did not improve significantly in the frequency of code 10, assessing whether claims or evidence were workable, valuable or convincing. More than half of the participants (n=31 out of 55, 56%) never used this skill pre-and-post-PBL (Figure 8.5), much higher than code 8 (n=18 out of 55, 33%) and code 9 (n=21 out of 55, 38%) in Evaluation skill (Figure 8.6).

Figure 8.5 The number of students wether used code 10 in writing





The slight improvement and a comparatively high percentage of code 10 suggested that students were not concerned or did not generate such an ability to evaluate the statements' feasibility or reliability in writing accurately. Therefore, it is worthwhile reminding students to explain whether their conclusions or perspectives are reasonable or some standpoints or behaviours are plausible when citing other people's words or making examples.

8.4.2.4 Overall CT skill development

A result of a Wilcoxon signed rank test (Table 8.12) reports that the P value of the sum frequencies of all ten codes (*p=.000, Cohen's d=.43*) is less than .05, indicating a significant difference in total frequencies (i.e. Inference+Analysis+ Evaluation) made. The effect size is .43, revealing a modest effect (Table 8.3).

Table 8.12 A Wilcoxon signed rank test of the frequencies of total thinking skills in writing during and after PBL

	all codes-post-PBL - all codes-pre-PBL
Z Asymp. Sig. (2-tailed)	-4.548[b] .000
D	.43

b. Based on negative ranks

The apparently increased sum of frequencies indicates that students enhanced their capacity to apply CT skills to their writing after using CT strategies in PBL classes. Meanwhile, this increased capacity for utilising thinking skills in their compositions further revealed that their mindset attribute of thinking critically generally increased.

8.5 Students' Perceptions of Good Essays: Results of the Interview

To further explore what might be the impact of CT-oriented PBL, 17 students were interviewed at the end of the teaching intervention programme. The key finding is that students' perception of a good essay has changed. Before PBL, students believed that linguistic features and handwriting were crucial to higher writing quality. However, after PBL, they regarded linguistic features and critical thinking as the priorities

8.5.1 Perceptions before PBL

Students concentrated on two strands when discussing writing standards in high school (before PBL). Eight (out of 17, 47%) students regard language as one criterion, including sentence patterns (n=5 out of 8, 63%), advanced vocabulary (n=4 out of 8, 50%) and grammar (n=4 out of 8, 50%). These three aspects typically merged when they came up with their evaluation (IS01 and IS10), indicating that students normally regarded them as integration when assessing the language in writing.

> (IS01): "I think a good essay should have advanced vocabulary and complex sentences. And the sentences were smooth."
>
> (IS10): "Well, I think sentence patterns, grammar and some advanced vocabulary was essential for a good essay."

In addition, seven (out of 17, 41%) students report that handwriting is the other standard (IS01 and IS07). The quotes showed that good handwriting was essential for higher writing scores in the examinations in high school. This standard is consistent with the research of Yu *et al*. (2021), pointing out that one criterion for assessing EFL writing is the quality of handwriting in the Chinese context.

> IS01): "I think the disparity between higher and lower score in writing was determined by good handwriting and mistakes of grammar points."
>
> (IS07): "Our teacher (in high school) told us that good handing writing was quite important. If we made it, we can get a better grade in writing."

High school students were usually given a specific writing direction with three points. Eleven (out of 17, 65%) students state that teachers explicitly remind them to write in three paragraphs following that direction. Therefore, they did not require to think about the structure or organisation. Furthermore, almost all interviewees mention that their teachers provide a template, especially for the first (the beginning section) and

the third paragraph (the conclusion section). Therefore, the writing direction and template made writing more accessible for students (IS07).

> (IS07): "We seldom deviated from the topic because we were given a writing direction. It told us what to write. Like, what is the question and what are your opinions. It's simple to follow the directions. It depended on whether you wrote good sentence patterns if you wanted to get a higher score."

Besides, what lecturers teach in high school is mainly concentrated on language points, for instance, vocabulary (n=16 out of 17, 94%), grammar (n=12 out of 17, 71%) and sentence patterns (n=10 out of 17, 59%). In comparison, fewer teachers teach thinking (n=5 out of 17, 29%) in writing, and the most frequently mentioned CT element is using linking words (n=4 out of 5, 80%). This finding echoes some Chinese scholars' perspectives (Huang, 1998; Li, 2011), who criticise the unidimensional English teaching priority of the linguistic accuracy and ignorance of the content students are trying to express. Compared to an explicit way of teaching linguistic knowledge, the way practitioners teach thinking is implicit because none mentions the term "critical thinking". The other teaching emphasis is handwriting (n=8 out of 17, 47%). Attaching the importance of language and neglecting thinking is not unique in China. Lee (2016) reports that secondary EFL classroom writing assessment is predominately language-focused in general, rather than the content and idea.

Results of the consistency of students' criteria and teachers' teaching contents in language and handwriting revealed that students were strongly influenced by their teachers, which was in accordance with the traditional value of honouring teachers and respecting teaching under Confucian culture. The inconsistency was students did not regard critical thinking as their standard, which can be attributed to teachers' implicit teaching method or neglect and the given writing directions in the examination.

8.5.2 Perceptions after PBL

When discussing the writing criteria in university (after PBL), students focused on two aspects. Firstly, fourteen (out of 17, 82%) students report that language is their important consideration, including vocabulary (n=10 out of 14, 71%), sentence patterns (n=9 out of 14, 64) and grammar (n=6 out of 14, 43%). This standard was the same as the one before PBL. Their unchanging linguistic standard with a higher number of students who regarded it as a criterion indicated language performance was the cornerstone when evaluating essays. In addition, thirteen (out of 17, 76%) students report that critical thinking is one of their criteria. Firstly, eight (out of 13, 62%) report that the organisation is their primary consideration (IS15). Secondly, five (out of 13, 38%) students state a clear thesis statement is essential (IS02). Finally, four (out of 13, 31%) students mentioned the necessity of a good structure and logic in writing (IS10 and IS18).

> (IS15): "I think the most important is a good organisation, which determined the readers' feeling. It will be hard for them to understand if the organization is chaotic."
>
> (IS02): "The primary standard for me was whether the thesis statement was proper and clear."
>
> (IS10): "Well, I think whether the structure was complete was essential. I mean whether we can state clear all the key information."
>
> (IS18): "I think the second important was logic in the essay."

The above quotes revealed that students began considering the organisation, structure, coherence and logic when evaluating writing. And these points are vital components of CT skills in writing. In addition, when being asked about the difference in standards to assess writing in high school and university, six (out of 17, 35%) students explicitly reported they needed to think deep, and five (out of 17, 29%) mentioned their writing should be more logical in university. This emerging criterion indicated that students went further to dig for deep thinking rather than staying at a superficial level, like handwriting. In this sense, they seemed to become aware of its importance and tended to think critically during the writing process.

Their transformed standard from dimensions of language and handwriting to language and CT echoes the finding of students' apparently increased CT skills in their compositions (section 8.4.2.4). Both qualitative data of interview and writing

validated each other and increased the validity of the results. Furthermore, both results revealed students' improved overall mindset towards CT. In addition, they supplemented the results of the CCTDI, which provided six dimensions of CT disposition without overall measurement.

8.5.3 Students' performance when doing peer review

Although students transformed their standards to evaluate the writing quality, what they most corrected was the linguistic mistakes (n=14 out of 17, 82%) on the one hand. The interviewees explained why language correction outperformed the thinking revision. At first, six (out of 14, 43%) students said it was difficult to apply the thinking rubric to revise an essay(IS04), while only two (out of 14, 14%) interviewees mentioned that they did not understand it. This quantitative contrast means the most significant challenge to revising thinking is not the difficulties in understanding the checklist but the application capacities.

> (IS04): "I think it's hard for me to use it (the rubric) to revise my writing. ... Maybe because I don't have much experience."
>
> (IS09): "I think it is much more difficult if I use it (the checklist) than revising by myself. And I think it is too general for me to understand."

In addition, three (out of 14, 21%) students forgot to use the rubric (IS07), and three (out of 14, 21%) thought it would cost them too much time to revise language and thinking at the same time (IS01). These two reasons indicated they did not place both aspects equally, but the language came first, especially when they did not have enough time. This finding is consistent with the study of Zhang *et al*. (2020), which attributes the main factor hindering Chinese teachers' CT education to the fact that students do not attach much importance to critical thinking on account of an obedient and passive learning habit.

> (IS07): "I put it (the checklist) in my textbook and forgot to use it."
>
> (IS01): "He (the writer) give us the essay on Monday, but we have to exchange our comments on Thursday. So we don't have enough time to revise."

Finally, two (out of 14, 14%) students treated peer review as an Error Correction, a question type in Gaokao requiring them to correct linguistic mistakes in a provided essay (IS17). It seemed that they still followed a familiar and easy way to revise language instead of thinking, even though they had changed their standards.

> (IS17): "I think the peer review in my group is like a Error Correction in Gaokao. What we concern are the same as it. Of course, we also notice the logical issues, but what we correct the most is language."

On the other hand, only six (out of 17, 35%) students stated that their groups used the thinking checklist, indicating its usage percentage was not high, which may affect their thinking evaluation quality.

In general, owing to students' comparatively poor application ability of the rubric of assessing CT, relatively inactive attitudes towards CT revision in writing, and pursuing a convenient way of merely revising language, students in this study laid higher stress on linguistic mistakes than the thinking problems. What they actually performed seemed inconsistent with their changed standards when evaluating writing quality. This inconsistency indicates that students increased their thinking awareness in writing. However, their thinking may be still at the stage of more perception and less behaviour.

8.6 Summary

This chapter explores the impact of CT-oriented PBL on students' thinking improvement in response to the first research question from three sources: CT mean scores from tests and standards of evaluating writing before and after PBL, and frequencies of using thinking skills of the first and final draft. Generally, their CT skills and disposition improved after PBL. Specifically, there are four following key findings in this chapter.

Firstly, the significantly different mean scores of CCTST suggested students developed overall and all the targeted core thinking skills of Inference, Evaluation and Analysis after PBL. Their overall CT ability stayed at a Strong level before and after PBL; the Inference enhanced from a Moderate to a Strong level; the Analysis and Evaluation remained at a Moderate level after PBL. Secondly, the significantly

increased frequencies of core thinking skills of Inference, Analysis and Evaluation in the essays indicated that students increased core thinking skills in writing. In addition, the increased total frequencies revealed that students developed significantly in their general CT skill. Thirdly, the changed mean scores in CCTDI revealed their CT dispositions developed unevenly: they did not develop their thinking dispositions towards Truth-seeking and Maturity in judgement, but Open-mindedness, Inquisitiveness, Analyticity and Confidence in Reasoning after PBL. Students held an Ambivalent attitude towards Open-mindedness and Truth-seeking, a Positive attitude towards Analyticity, Confidence in reasoning and Maturity in judgement and a Strong Positive attitude towards Inquisitiveness. Among all the dimensions, students held the most negative attitudes towards Truth-seeking with the lowest mean score, while the most positive attitudes towards Inquisitiveness with the highest mean score before and after PBL. Finally, students transformed writing standards from linguistic features and handwriting to linguistic features and CT, indicating an enhanced CT consciousness in their writing.

Chapter 9 Findings: EFL Students' Attitudes towards CT-oriented PBL Intervention

9.1 Introduction

This chapter reports findings in response to the second research question, which investigates students' attitudes towards the effectiveness of CT-oriented PBL pedadogy to improve their CT. Firstly, it presents the results of a CT strategies questionnaire, including mean scores of students' attitudes towards the effectiveness of CT strategies in enhancing thinking skills (Part I of the questionnaire) and dispositions (Part II) and attitudes to each CT strategy (Part III). The results of the Friedman tests present whether the above results are significantly different and which one was the most effective strategy. Secondly, it illustrates the findings of a PBL questionnaire, including mean scores and results of the Friedman test are subsequently provided to test whether they are statistically significant. Finally, it reports the results of interviews about students' attitudes towards and perspectives of using three CT strategies.

As discussed in Chapter Five, the quantitative and qualitative data aimed to achieve cross-validation. Students' attitudes in this study were evaluated from two sources: the quantitative data from two questionnaires and qualitative data from interviews (Figure 5.2) after PBL intervention. On the one hand, quantitative results from questionnaires (i.e. CT strategies and PBL questionnaire) through descriptive statistics (i.e. mean scores) illustrate students' attitudes towards CT strategies and PBL teaching intervention. At the same time, inferential statistics of the Friedman test examine whether mean values of attitudes are significantly different among three core thinking skills (i.e. Inference, Analysis and Evaluation), six dimensions of CT disposition (Open-mindedness, Inquisitiveness, etc.) and three elements in PBL (i.e. teachers' role, students' awareness of the real world and their work in group). If they did, it indicates the values of their attitudes in this study have statistical significance (Dörnyei and Taguchi, 2010). Furthermore, post hoc tests (i.e. the Wilcoxon signed rank tests) explore which specific skill and mindset attribute students made the greatest or the slightest achievement or which PBL element students believed helped them improve CT the most. On the other hand, the qualitative results from students' individual

face-to-face interviews through descriptive statistics validate the findings of the questionnaires.

9.2 Results of the CT Strategies Questionnaire

The questionnaire was designed to examine students' attitudes towards the effectiveness of CT strategies to improve their generic critical thinking (Part I+II, 15 items) in two dimensions: CT skills (Part I, items 1-6) and CT dispositions (Part II, items 7-15) (see Chapter 6, section 6.3.1, Table 6.2). In addition, it examined participants' attitudes towards the effectiveness of each CT strategy in improving students' CT (Part III, items 16-18) (Table 6.2). All 18 items were in the form of multiple-choice questions with six-point Likert scales from strongly disagree (1 point) to strongly agree (6 points). Generally, if the mean value is equal to or higher than 4.0, it represents positive attitudes; if lower than 4.0, it represents negative attitudes (see section 6.5.2.1). Specifically, the mean score of 4.00-4.99 represents students' essential positive attitude; the 5.00-5.99 point represents their positive attitude, and the 6.00 point represents their strong positive attitude. In contrast, the mean score of 1.00-1.99 shows students' strong negative attitudes; 2.00-2.99 indicates a negative attitude, and 3.00-3.99 shows their slightly negative attitude. Descriptive statistics and the results of the Friedman test are presented for each section of the questionnaire.

9.2.1 Descriptive statistics: students' attitudes the effectiveness of CT strategies

Table 9.1 below presents the means (M) and standard deviations (SD) of the values of participants' attitudes (Part I and II of the questionnaire). The mean value of all the core thinking skills is 5.09 (Table 9.1, Part I), indicating students held a positive attitude towards the effectiveness of CT strategies to improve the generic CT skill. Specifically, the mean values for three core skills are above 4.00: Evaluation (*M=5.18, SD=.776*), Analysis (*M=5.14, SD=.852*) and Inference (*M=4.94, SD=.904*). These results indicated that students held a positive attitude toward CT strategies to improve the Evaluation and Analysis and an essential positive attitude to Inference.

Table 9.1 Descriptive statistics of participants' attitudes to the effectiveness of CT Strategies to improve CT (Part I and II of the CT strategies Questionnaire)

Questionnaire Sections	N of items	Mean score	Mean score for each sub-skill/ sub-scale		SD
Part I: The attitudes towards the CT strategies to improve CT skills	6	5.09	Evaluation	5.18	.776
			Analysis	5.14	.852
			Inference	4.94	.904
Part II: The attitudes towards the CT strategies to improve CT dispositions	9	4.96	Open-mindedness	5.18	.816
			Inquisitiveness	5.00	.939
			Maturity in judgement	4.96	.903
			Confidence in reasoning	4.95	.834
			Truth-seeking	4.94	.827
			Analyticity	4.89	.915
The Mean of 15 items = 5.01					

In addition, the mean score of 4.96 of all the six dimensions indicated that students held a positive attitude towards CT strategies' effectiveness in improving their critical dispositions (Table 9.1, Part II). Specifically, mean values for four mindset attributes of Truth-seeking (*M=4.94, SD=.827*), Analyticity (*M=4.89, SD=.915*), Maturity in judgement (*M=4.96, SD=.903*) and Confidence in reasoning (*M=4.95, SD=.834*) are at the range of 4.00-4.99 (Table 9.1, Part II), indicating an essential positive attitude. In addition, mean scores of Open-mindedness (*M=5.18, SD=.816*) and Inquisitiveness (*M=5.00, SD=.939*) are higher than 5.00, indicating that students seemed to hold more positive attitudes than the other four mindsets. Finally, the mean value of all the 15 items (i.e. Part I of CT skills and Part II of CT disposition) is 5.01, suggesting that students held a positive attitude towards the effectiveness of CT strategies to improve their CT.

Besides the investigation of students' positive attitude towards the effectiveness of three CT strategies as integrity, it also examines their attitudes towards the effectiveness of each CT strategy (Part III of the questionniare). The mean value of the thinking maps strategy (*M=5.03, SD=.809*) (Table 9.2) indicates a positive attitude within the 5.00-5.99 range. The other two strategies of group discussion (*M=4.80, SD=.862*) and peer review (*M=4.78, SD=.894*) are in the range of 4.00-4.99, revealing an essential positive attitude.

Table 9.2 Descriptive Statistics of participants' attitudes to the effectiveness of three CT strategies (Part III of the questionnaire)

Questionnaire Sections	N of items	Mean score	Mean score for each sub-skill/ sub-scale		SD
The attitudes towards the CT strategies	3	4.87	Thinking maps	5.03	.809
			Group discussion	4.80	.862
			Peer review	4.78	.894

9.2.2 The Friedman tests for investigating students attitudes

9.2.2.1 CT skills development

Distribution analysis

Distribution analysis of was conducted first to examine whether data was distributed normally before conducting a repeated ANOVA (normal distribution) or a Friedman test (abnormal distribution) to examine whether students' attitudes towards each core thinking skill differed significantly (part I of the questionnaire). Sig. values of Kolmogorov-Smirnov (*p=.000*) and Shapiro-Wilk tests (*p=.000*) (Table 9.3) are less than .05, indicating a violation of the assumption of normality (Pallant, 2016). Therefore, the values are not distributed normally.

Table 9.3 Distribution of the results of three CT skills

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Evaluation	.233	94	.000	.867	94	.000
Analysis	.182	94	.000	.898	94	.000
Inference	.137	94	.000	.930	94	.000

Results of a Friedman test: attitudes towards the CT skills' development

The Friedman test results indicated a statistical difference among the mean ranks of three core skills with a less than .05 Sig. value ($\chi 2(2, n=94) = 17.64$, $p=.000$) (Table 9.5). Inspection of the median score for each skill is 5.00 (Table 9.4), indicating positive attitudes.

Table 9.4 Descriptive statistics and ranks of three core thinking skills

		Percentiles			Ranks
	N	25th	50th (Median)	75th	Mean Rank
Evaluation	94	5.00	5.00	5.50	2.14

Analysis	94	5.00	5.00	5.50	2.13
Inference	94	4.50	5.00	5.50	1.72

Table 9.5 A statistical difference in the mean rank of CT strategies

Test Statistics[a]

N	94
Chi-Square	17.643
Df	2
Asymp. Sig.	.000

a. Friedman Test

Having established a statistically significant difference among CT skills, post-hoc tests were followed up to examine where the differences actually occurred. It involved separate Wilcoxon signed rank tests (using a Bonferroni adjusted alpha value) (Pallant, 2016, p.242) on three pairs: Inference to Evaluation, Inference to Analysis and Analysis to Evaluation. It involved three tests; therefore, a revised alpha level for determining statistical significance would be .05 (in this study) divided by 3=.017. In other words, if the p value was smaller than .017, each pair had a significant difference. At this level of significance, post hoc analysis with Wilcoxon signed rank tests displayed a significant difference between Inference and Evaluation (*Z= -3.112, p=.002*) and Inference and Analysis (*Z= -3.166, p=.002*) (Table 8.5), indicating that the value of Inference was apparently different from the other two skills. Combined with its lowest mean score (section 9.2.1.1, Table 9.1), students were inclined to believe that they made the slightest improvement in inference. However, no significant difference was found between Analysis and Evaluation (*Z= -.505, p=.614*) (Table 9.6), indicating that students thought they made the same degree of improvement in the skills of Analysis and Evaluation by using CT strategies.

Table 9.6 Statistics of the Wilcoxon signed rank test of three core thinking skills

Test Statistics[a]

	Analysis - Evaluation	Inference - Evaluation	Inference - Analysis
Z	-.505[b]	-3.112[b]	-3.166[b]
Asymp. Sig. (2-tailed)	.614	.002	.002

a. Wilcoxon Signed Ranks Test b. Based on positive ranks.

9.2.2.2 CT disposition development

Distribution analysis

Following the same reason elaborated in the preceding section, the normal distribution of the values of part II of the questionnaire was examined. Sig. values of Kolmogorov-Smirnov and the Shapiro-Wilk tests were <.05, which indicated that the scores were not normally distributed (Table 9.7, see below) (Pallant, 2016).

Distribution analysis was conducted first to examine whether data was distributed normally before conducting a repeated ANOVA (normal distribution) or a Friedman test (abnormal distribution) in order to examine whether students' attitudes towards developing six CT dispositions differed significantly. Sig. values of Kolmogorov-Smirnov (*p=.000*) and Shapiro-Wilk tests (*p=.000*) (Table 9.7) are less than .05, indicating a violation of the assumption of normality (Pallant, 2016). Therefore, the values are not distributed normally.

Table 9.7 Distribution of the results of six CT dispositions

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Truthseeking	.307	94	.000	.809	94	.000
Analyticity	.116	94	.003	.963	94	.010
CTconfidence	.281	94	.000	.840	94	.000
Openmindedness	.295	94	.000	.737	94	.000
Inquisitiveness	.234	94	.000	.838	94	.000
Maturity in judgement	.327	94	.000	.778	94	.000

a. Lilliefors Significance Correction

Results of a Friedman test: CT disposition

Results of the Friedman test revealed a statistical difference in the mean ranks of six dimensions with a less than .05 ($\chi 2$ (2, n=94) = 22.68, *p*=.000) (Table 9.9). Inspection of the median value for each dimension is 5.0 (Table 9.8), revealing students held positive attitudes.

Table 9.8 Descriptive statistics and ranks of six CT mindset attributes

	N	Percentiles			Ranks
		25th	50th (Median)	75th	Mean Rank
Truth-seeking	94	5.00	5.00	5.00	3.40
Analyticity	94	4.50	5.00	5.25	3.02
CT confidence	94	4.75	5.00	6.00	3.43
Open-mindedness	94	5.00	5.00	6.00	4.04
Inquisitiveness	94	4.00	5.00	6.00	3.52
Maturity in judgement	94	5.00	5.00	5.25	3.60

Table 9.9 A statistical difference in the mean rank of CT disposition

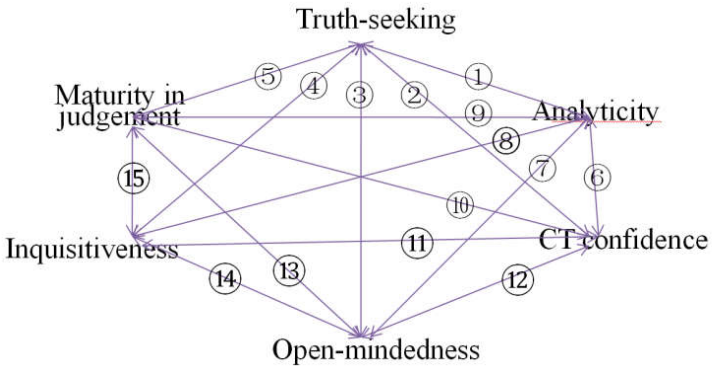
Test Statistics[a]

N	94
Chi-Square	22.676
Df	2
Asymp. Sig.	.000

a. Friedman Test

Having established a statistically significant difference among CT dispositions, post-hoc tests were followed up to examine where the differences actually occurred. It involved separate Wilcoxon signed rank tests (using a Bonferroni adjusted alpha value) (Pallant, 2016, p.242) on 15 pairs of 6 mindset attributes (Figure 9.1). For example, the first pair was Truth-seeking to Analyticity marked ①, and the second pair was Truth-seeking and CT confidence marked ②.

Figure 9.1 15 combinations of dimensions of CT disposition



In this sense, the post-hoc tests involved 15 separate Wilcoxon signed rank tests. Following the calculation method discussed in 9.2.2.1, a revised alpha level for determining statistical significance would be .05 (in this study) divided by 6=.008. In other words, if the *p* value was smaller than .008, there was a significant difference

among each pair. At this level of significance, post hoc analysis with Wilcoxon signed rank tests displayed a significant difference between only two pairs of Truth-seeking and Open-mindedness (*Z= -2.71, p=.007*) and Analyticity and Open-mindedness (*Z= -3.94, p=.000*) (Table 9.10).

Table 9.10 Statistics of the Wilcoxon signed rank test of the CT disposition

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮
Z	-.92[b]	-.01[c]	-2.71[c]	-.49[c]	-.34[c]	-.65[c]	-3.94[c]	-1.48[c]	-1.31[c]	-.61[c]	-.66[c]	-2.51[c]	-2.09[b]	-2.34[b]	-.046[b]
Asymp. Sig. (2-tailed)	.359	.994	.007	.626	.732	.517	.000	.140	.190	.545	.508	.012	.036	.019	.964

b. Based on positive ranks. c. Based on negative ranks.

Since no single mindset attribute differed from the other five, there was no statistical difference among the six dimensions of CT disposition, indicating students thought they made the same development in these six mindset attributes.

9.2.2.3 Three CT strategies

Distribution analysis

Following the same reason elaborated in the preceding section, the normal distribution of the values of part III of the questionnaire was examined. Sig. values of Kolmogorov-Smirnov and the Shapiro-Wilk tests of three CT strategies are <.05, indicating that the scores were not normally distributed (Table 9.11) (Pallant, 2016).

Table 9.11 Test of normality of three CT strategies

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Thinking maps	.314	94	.000	.765	94	.000
Group discussion	.295	94	.000	.841	94	.000
Peer review	.311	94	.000	.824	94	.000

The results of the Friedman tests: three CT strategies

The results of Friedman tests in Table 9.12 illustrated that the median value was 5.0, indicating that students held positive attitudes towards these three strategies. There was a statistical difference in the mean rank of three CT strategies with a less than .05 Sig. value ($\chi^2(2, n=94) = 12.94, p=.002$) (Table 9.13).

Table 9.12 Descriptive statistics and ranks of three CT strategies

	N	Percentiles			Ranks
		25th	50th (Median)	75th	Mean rank
Thinking maps	94	5.00	5.00	6.00	2.19
Group discussion	94	4.00	5.00	5.00	1.91
Peer review	94	4.00	5.00	5.00	1.90

Table 9.13 A Friedman test of the three CT strategies

N	94
Chi-Square	12.944
Df	2
Asymp. Sig.	.002

Post hoc tests by the Wilcoxon signed rank test displayed a significant difference between pairs of thinking maps and group discussion (Z= *-2.838,* p=*.005*) and between thinking maps and peer review (Z= *-2.948,* p=*.003*). However, there was no significant difference between group discussion and peer review (Z= -.393, p=.694) (Table 9.14). Therefore, the value of thinking maps was obviously different from the other two CT strategies. Combined with the result of its highest mean score (see section 9.2.1, Table 9.2), students believed that the thinking map strategy was the most effective in improving CT. This result differs from Wang and Seepho 's (2017) study, showing the strategy of group discussion is better than the thinking map, and both strategies significantly help students think critically.

Table 9.14 Statistics of Wilcoxon signed rank tests of the three CT strategies

	Group discussion - Thinking maps	Peer review - Thinking maps	Peer review - Group discussion
Z	-2.838[b]	-2.948[b]	-.393[b]
Asymp. Sig. (2-tailed)	.005	.003	.694

b. Based on positive ranks.

9.3 Results of the PBL Questionnaire

This section reports participants' attitudes towards the effectiveness of three key components of PBL class to enhance CT (see section 6.3.1, Table 6.3): teacher's role (items 1 and 2), students' group work (items 3 and 4) and students' awareness of the real world (item 5). In addition, students' general attitudes to the PBL class were also explored (item 6). As discussed in section 9.2, this questionnaire followed the same format and scoring method as the CT strategies questionnaire. In addition, one open

question of asking their suggestions for PBL. Descriptive statistics and the results of the Friedman were illustrated orderly.

9.3.1 Descriptive statistics: students' attitudes towards the effectiveness of PBL

Table 9.15 displays descriptive statistics of the mean values (M) and standard deviation (SD) of students' attitudes towards the effectiveness of PBL. The mean value of their attitude towards PBL is 4.78 at 4.00-4.99, indicating an essential positive attitude. Among the three PBL crucial components, the mean values of 5.19 and 5.04 indicated that students held positive attitudes towards a teacher's scaffolding role and the awareness of the natural world and the essential positive attitudes towards students' group work with a mean score of 4.98. These descriptive statistics revealed that students held positive attitudes towards the effectiveness of PBL intervention in general and towards all the three components to improve CT.

Table 9.15 PBL questionnaire: descriptive statistics

Questionnaire Sections	N of items	Mean score	Mean score for students' attitudes towards	
Attitudes towards different elements in PBL	5	5.08	Teacher's scaffolding role in PBL	5.19
			Awareness of the real world when doing project work	5.04
			Group work	4.98
Attitudes towards the effectiveness of PBL	1	4.78	Effectiveness of improving CT	4.78

The Mean of 6 items = 5.03

9.3.2 The Friedman test for investigating which PBL component students held the most positive attitude

Distribution analysis

Following the same reason as mentioned previously (see 9.2.2.1), the values were examined to whether they were normally distributed. Table 9.16 below displayed the Sig. values of Kolmogorov-Smirnov and the Shapiro-Wilk tests were .000 (<.005), indicating a violation of the assumption of normality (Pallant, 2016), so the values were distributed abnormally, and a Friedman test was conducted to test whether the scores were obviously significant.

Table 9.16 Distributions of three PBL components

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
A teacher's role	.188	94	.000	.908	94	.000
Group work	.220	94	.000	.830	94	.000
Attention on the real world	.277	94	.000	.822	94	.000

a. Lilliefors Significance Correction

Results of the Friedman test

Results of the Friedman test showed that the median value was 5.0 for each, and the percentiles indicated that students generally tended to be positive (Table 9.17). There was no statistical difference in the mean rank of PBL components with a more than .05 Sig value ($\chi^2(2, n=94) = 3.25$, $p=.197$) (Table 9.18), indicating an insignificant difference between at least two of them. In other words, students held the same positive attitudes towards these three components to help them improve their thinking.

Table 9.17 Descriptive statistics and ranks of three PBL components

		Percentiles			Ranks
	N	25th	50th (Median)	75th	Mean rank
A teachers' role	94	5.00	5.00	5.50	2.12
Group work	94	4.50	5.00	5.50	1.92
Attention on the real world	94	5.00	5.00	6.00	1.96

Table 9.18 A Friedman test of three PBL components

N	94
Chi-Square	3.246
Df	2
Asymp. Sig.	.197

In this light, this statistically non-significant result indicated no difference in students' attitudes towards three PBL components (i.e. teachers' scaffolding role, awareness of the real world and group work). In other words, students held the same extent of positive attitudes towards the effectiveness.

9.3.3 Content analysis of the open question: suggestions for PBL teaching

Following the analytical steps in section 6.5.2.2, a content analysis method analysed 88 English transcripts from 72 respondents (marked QS01-QS72) with 1048 English

words. Because one student may provide more than more suggestions, several suggestions for the same person were marked .1 or .2. For example, if I wanted to cite the second suggestion of the No. 01 student, I marked QS01.1. If the respondent only wrote one suggestion, I also marked it as .1, in accordance with the format.

The content analysis results of the open-ended question presented three aspects of students' suggestions. Most of them mentioned CT strategies (n=27 out of 88, 30.7%). Twenty-one (out of 27, 77.8%) students advocated continuing to use these CT strategies. Some respondents simply suggested they could use them for the next semester (n=5 out of 21, 23.8%), or they should reflect on how to make these strategies more effective (n=3 out of 21, 14.3%). More of them (n=13 out of 21, 61.9%) went a step further and provided specific suggestions for teachers and students to revise CT teaching or learning. On the one hand, they suggested students could distribute collaboration when discussing in groups (QS22.2), write down their reflections of the discussion (QS24.2) or draw a thinking map of the passage in the textbook instead of teachers' modelling (QS43.2). On the other hand, teachers should increase discussion time (QS23.2, QS41.2) and select and analyse model essays as part of the lecture (QS43.1). In contrast, 6 (out of 27, 22.2%) students suggested that they should reduce to use CT strategies, such as decreasing discussion time (n=5 out of 6, 83.3%) and drawing thinking maps (n=1 out of 6, 16.7%). The highest amount of students' suggestions of thinking teaching indicated that most participants endorsed the effectiveness of CT strategies, which was in line with students' positive attitudes towards the CT-oriented PBL (Section 9.2).

The second highest frequencies of suggestions (n=20 out of 88, 22.7%) were about adding CET4 teaching contents in class. In comparison with vaguely mentioning that teachers should teach them to prepare for CET4 (n=5 out of 20, 25%), most of them clarified that they wanted teachers to provide professional training on different question types of that test (n=13 out of 65%) or lecture some specific skills to get higher grades (n=2 out of 10%). This second highest frequency indicated their high demand for CET4 training, even though they had at least one year to participate in the tests. As one of their priorities in English learning, it was supposed to be related to the assessment-oriented learning context, a factor hindering students' thinking development (chapter 11, section 11.2.3.1).

The third strand was their hope for more lectures about linguistic knowledge (n=18 out of 88, 20.5%). They seemed to have a clear direction for acquiring related knowledge, such as vocabulary or sentence patterns (n=8 out of 18, 44.4%), writing (n=6 out of 18, 33.3%) and grammar (n=4 out of 18, 22.2%). Their demand for language learning was the same as the knowledge they acquired in high school, except the writing. In a nutshell, students' suggestions for thinking strategies and language learning were consistent with their double standards of evaluating writing (section 8.5.2). Besides, their learning demand for preparing for the test indicated that the exam-oriented context might influence their English study at university.

9.4 Results of the Interview

This section reports the results of the 8th open question of the interview: their attitude towards the effectiveness and preference of three CT strategies. Most interviewees believe thinking maps improve their CT (n=15 out of 17, 88%). The strategy of group discussion comes second (n=10 out of 17, 59%). And peer review employs the least number (n=7 out of 17, 41%). This finding supports the result of the CT strategies questionnaire in section 9.2.2.3, reporting that students held the most positive attitude towards the effectiveness of thinking maps. Their benefits in improving students' CT will be discussed in chapter 10.

Although peer review has the least number of students who recognise its effectiveness, it employs the largest amount of students (n=13 out of 17, 76%) who prefer this strategy, which is far beyond the number of thinking maps (n=3 out of 17, 18%) and group discussions (n=1 out of 17, 0.6%) in this study. Thus, a gap exists between a great extent of their liking and a comparatively low utility to improve critical writing performance. This gap revealed that students had great potential for utilising this strategy to enhance their thinking skills in writing. Most of them showed enthusiasm for this strategy, indicating that they may employ high motivation. And this motivation is closely related to better learning outcomes (Wang, B., 2013; Zhai, 2021). In this sense, if this gap failed to bridge, students were more likely to reduce their impetus in the long run. Therefore, exploring factors hindering the effectiveness was indispensable, and detailed information will be discussed in sections 10.3.2, 10.4.2 and 10.5.2.

9.5 Summary

This chapter reported students' attitudes towards the effectiveness of CT strategies and PBL intervention to enhance their CT from two questionnaires and interviews. Generally speaking, students held a positive attitude towards the effectiveness of CT strategies and PBL teaching intervention. Specifically, they believed that their Inference skills improved the least, and their six CT mindset attributes developed the same. In addition, the strategy of the thinking map was the most effective. Finally, three PBL key elements played an equal role in developing their CT.

Chapter 10: Findings: How CT Strategies Might Affect EFL Students' Thinking in PBL Classes

10.1 Introduction

This chapter presents findings in response to the third research question of how students developed thinking skills in CT-oriented PBL classes. As discussed in chapter five, the interview aimed to explore students' perceptions of how three CT strategies (i.e. thinking maps, group discussion and peer review) developed students' thinking. As discussed in the coding list in Chapter 6, section 6.5.4.1, the last three interview questions were divided into two categories: (1) the specific CT skills students developed after using each CT strategy, and (2) the issues they were concerned about (i.e. challenges or problems) when using the strategies. Following the analytical procedures mentioned in Section 6.5.4.2, a content analysis method analysed English transcripts of 17 interviewees (marked IS01-IS17) with 57,336 English words.

10.2 Participants

The 17 interviewees were students from the stomatological class in the School of Stomatology, Binchou Medical University, joining the CE course as part of a five-year diploma program. Their English levels were supposed to be at level four according to the requirement of CSE (2018) and the equivalent of point 4.5 in IELTS (Dunlea *et al.*, 2019, p.113). They were expected to reach level five (NEEA, 2018), an equivalent of point 5.5 in IELTS (Dunlea *et al.*, 2019, p.113) when finishing their two semesters of English study. Table 10.1 displays all the interviewees' background information. Their names were anonymous in different marks according to their order of being interviewed. Taking *IS01* for example, *IS* represented students who took part in the interview and *01* represented the first student who accepted the interview.

Table 10.1 Interviewees' background information

Participants	Sex	Group and class	English score in Gaokao (out of 150)
1. IS01	Female	Group 2, class 2	124
2. IS02	Female	Group 3, class 2	130
3. IS03	Male	Group 2, class 1	123
4. IS04	Female	Group 8, class 2	126
5. IS05	Male	Group 3, class 1	112

6. IS06	Female	Group 6, class 1	130
7. IS07	Male	Group 4, class 1	125
8. IS08	Female	Group 8, class 1	135
9. IS09	Male	Group 1, class 1	134
10. IS10	Female	Group 5, class 1	119
11. IS11	Male	Group 7, class 1	120
12. IS12	Female	Group 7, class 2	111
13. IS13	Female	Group 10, class 1	133
14. IS14	Female	Group 9, class 2	130
15. IS15	Male	Group 10, class 2	122
16. IS16	Female	Group 4, class 2	126
17. IS17	Male	Group 9, class 1	123

Since there was no official document linking the grade in Gaokao to IELTS, their scoring average was a standard to evaluate their English level. In this study, the interviewers' scoring average was 83.2%, and their average English level was at least medium-high.

Their English learning experiences were stated in the following section:

Case one (IS01) thought the writing skills she learnt in Chinese class, like comparison and contrast, could be transferred to her English writing because her English teacher in high school seldom taught such knowledge. The key element that led to her success in Gaokao was vocabulary. She did not seem accustomed to the current CT-centred English learning method because she was heavily influenced by her previous experience focusing on vocabulary and grammar learning.

Case two's (IS02) attitude towards previous English learning was a paradox. On the one hand, she was not confident in her competence to improve English and thought teachers should play a central role in class. On the other hand, she complained about teachers' tedious teaching methods, mainly focusing on linguistic knowledge.

Case three's (IS03) English learning was test-oriented, and memorising model essays was his primary way of learning English writing.

Case four (IS04) came from a small town and studied English to get higher marks, so he did a lot of practice. His teacher in high school focused on handwriting and good sentence patterns at the beginning and end; therefore, he never experienced thinking

training. Deeply influenced by didactic teaching, he preferred a teacher-dominated class.

Case five (IS05) excelled at maths, physics and chemistry, but not at English. He only started revising English 100 days before Gaokao. Apart from learning at school, he also had one-to-one tutoring to learn tricks and tips to pass the exam, including a template for improved writing performance. As a result, he increased his writing scores from 5 or 6 points to more than 20 out of 25 points. Besides the tricks and a template, he believed more writing practice before examination helped him perform well in writing in Gaokao. Therefore, although his score (112 points) was much lower than the average level (125 points) among the interviewees, he felt satisfied with the outcome.

Case six's (IS06) former English teacher often taught writing and paid great attention to the linking words besides advanced sentence patterns. She usually spent nearly half of her exam time on writing. She felt teachers played a different role in English teaching in university. While teachers in high school were regarded as supervisors pushing her to study, teachers in university were regarded as assistants assisting her in learning more efficiently. She held a positive attitude towards both roles.

Case seven (IS07) noticed huge differences when comparing his English learning experiences in high school with the university. He learned writing by memorising sentence patterns in high school regarding essay writing. While in university, they learnt to analyse the structure and learn writing skills of the passages. The teaching method in high school was more passive. In contrast, it was more interactive at the university.

Case eight (IS08) did a great job as an English class representative whose duty was to assist the teacher in reminding students to complete homework, forward teachers' messages to her classmates, and so on. Although her score in the examination was comparatively high, she reflected she needed to improve her application skills instead of obtaining a high score on tests. Instead of merely passing the CET4, she had high expectations of English study to improve her speaking and writing.

Case nine (IS09) contributed his success to his experienced English teacher in high school, who helped him improve his score by focusing on vocabulary and grammar.

To some degree, he was the beneficiary of teacher-centred teaching. Compared to his previous learning experience, he felt that what he learned in the current class was applying English in practice rather than getting a higher score. However, he thought it hard to achieve a satisfactory result. He attributed this difficulty in acquiring a higher practical capacity for using English to his insufficient vocabulary. He also found it hard to memorise new vocabulary without teachers' strict supervision. He hoped teachers to play a more significant role in providing standard answers to their discussion or thinking maps.

Case ten (S10) did not achieve the average score among these 17 interviewees (125 points). However, she claimed she underperformed in Gaokao because she usually got at least 130 or even 140 points in mocks. She emphasised that application was critical, and she even thought it was meaningless to learn English if she couldn't use it in reading or writing. Therefore, she advocated learning English by writing instead of simply memorising vocabulary and sentence patterns.

Case eleven (IS11) also did not achieve the average score among the interviewees. His motivation to learn English was driven by his strong desire to get the highest score in Gaokao. Under the high pressure in high school, he wanted to relax for a while after entering university. Therefore, he lost motivation to study, including English. He believed learning vocabulary and grammar was vital to preparing for the coming CET4. He also thought passing that exam was far more crucial than developing English language skills.

Case twelve (IS12) achieved the lowest score (111 points) among all the interviewees and ranked fourth from the bottom in her class (n=50). Like most interviewees in this study, her learning experiences in high school were teacher-dominated and focused on vocabulary and grammar. In addition, she reflected that she lacked the self-discipline to spend more time studying after class because no teacher forced her to learn anymore. At the same time, she greatly benefited from peer review because she thought her groupmates' English was much better, and they always encouraged her.

Case thirteen (IS13) was the only one who claimed that group discussion was an activity in her English class in high school. Besides the traditional extensive reading class, she also had a separate English writing and public speaking class. She favoured

this collaborative learning by saying it benefited her cooperative, problem-solving, self-study and independent thinking abilities except for linguistic development. In addition, her collaborative learning experiences and enrichment courses made her more confident when answering teachers' questions and discussing in groups.

Case fourteen (IS14) placed English learning and specialised courses (i.e. Anatomy) were her priority. She thought interaction among students became stronger in current classes, and her group even set up an online chat group to share literature for writing. Their communication in class and after class increased their familiarity, which was essential for these newcomers to the university.

Case fifteen (IS15) described his previous English learning experiences in class as dull and drowsy. No matter how hard he studied, his score was always between 120 and 130 points, so he was less motivated to learn English and not interested in it. His negative attitude in high school seemed to transform after entering the university due to his future academic plan to join the master' program with a requirement of high English level. In other words, he was motivated to study harder to join a further study. Although having a strong desire to learn hard, he did not engage in learning. For example, he said that he did not take notes about the results of peer review, did not prepare for the group discussion and complained that no one could lead him to study after class.

Case sixteen (IS16) believed that the key to English learning was memorising vocabulary, and her learning experience also followed the traditional teacher-centred method. Her group was the only group with a clear work allocation. For example, someone took notes, and someone else took responsibility for searching for examples to support the perspectives. This collaborative learning partly released her pressure to write more than a 500 words essay because she had never done it before.

Case seventeen (IS17) had the experience of studying in groups monitoring each other to complete homework by following teachers' systematic schedules in high school. Therefore, he was accustomed to group study quickly. At the same time, he thought he needed more self-discipline to make a schedule for academic study because no one pushed them to study anymore.

10.3 How Thinking Maps Developed Students' Thinking

Several benefits of thinking maps and corresponding core thinking skills development were reported first. Then, it revealed several issues students were concerned about when drawing maps.

10.3.1 The core thinking skills deveopment by using thinking maps

The interview results revealed that students thought thinking maps helped them improve three skills: Interpretation, Analysis and Inference. The interpretation means providing a framework, and discovering and describing the content and meaning (Chapter 2, section 2.2.4, Figure 2.1). Eleven (out of 17, 65%) students said that thinking maps helped them have a clear mind about what to write (IS06 and IS16), resulting in their deep understanding of the logical relationships and a more comprehensive structure for their essays.

> (IS06): "My mind will be at a mess if I don't have a thinking map before writing. I don't know what to write. Vice versa, I know what to write in each paragraph because I have a framework."
>
> (IS16): "I am clear from which aspects could I write when elaborating an issue, and I will write them all one by one."

Their quotes suggested that thinking maps visualise all the key ideas, allowing students to list, determine, and categorise them according to the branches. In addition, the hierarchies of ideas and examples help explain or support claims based on the sub-branches in practice. In this sense, contents in thinking maps served as a holistic guideline to remind students what they should write. Otherwise, their writing may be stuck on the grounds of lack of focus, logical thinking, or coherence (see IS08).

> (IS08): "It's like you stacked several sentences together and suddenly you don't know what should you add to those sentences."

When following thinking maps, students were disposed to precisely comprehend and describe the thesis statements and main claims and then transform concepts of the maps into words in the essays. In this way, students developed the Interpretation skill. This result supported the previous findings of the thinking maps' effectiveness in

producing and classifying valid ideas, interpreting information preciously, generating supporting details, and comprehending arguments (Al-Shaer, 2014; Daley and Torre, 2010; Kaeppel, 2021; Lee, 2013; Liu, 2011).

Analysis refers to examining ideas of relationship identification between the part and the whole and detecting arguments, including assertions and corresponding justifications (Figure 2.1). Eight (out of 17, 47%) students reported that thinking maps made their writing more organised. For example, their thinking became clear and logical (IS01, see below).

> (IS01): "... I wrote the composition directly without a thinking map due to the limited time when taking part in the examinations. I think it looked like a patchwork. But, I felt my writing was clear and coherent if I drew the thinking map. I knew exactly what to write firstly, secondly and thirdly."
>
> Interviewer: "What do you mean by clear?"
>
> (IS01): "The organisation and logic."

The orderly organisation manifested in students' increased awareness of writing sequence (IS07, see below) and their enhanced capacity to determine writing focus according to the extent of relevance to the topic in an inferential relationship (IS14, see below).

> (IS07): "... The structure became clearer. Like, I knew clearly about what should I write next."
>
> (IS14): "I would state my ideas according to different branches in the thinking map. At first, I considered which branches were the key and which ones were less critical. I wrote everything I could for the key branches to emphasize and reduce the words of the less important branches. ... For example, I would elaborate more on the reasons and influences in detail and write the background information less because it was less important."

The above quotes indicated that drawing thinking maps made them more likely to produce a logical organisation. This finding was consistent with Rahman and Ambreen's study in 2018, reporting that thinking maps enable the essays to be more organised and systematic.

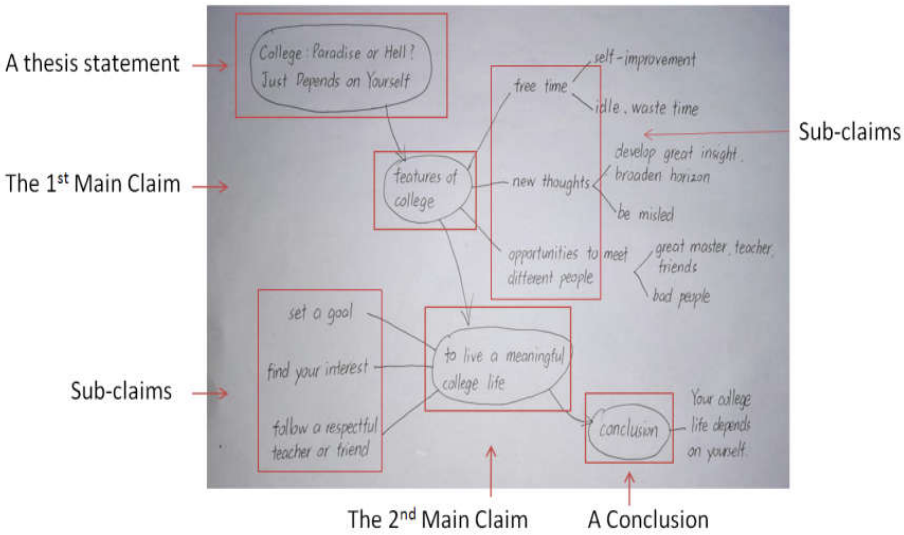
The other manifestation of organisation resided in their consciousness of coherence in writing. Four (out of 17, 24%) students mentioned that thinking maps reminded them

to focus on the thesis statement and not miss any key points (IS03 and IS06, see below).

> (IS03): "It reminded me not missing the content related to the topic because I wrote all the key points in that map."
>
> (IS06): "I used to write whatever came to my mind and the content was messy sometimes. While my essay became better by the thinking map because it restricted my thinking to avoid straying."

The quotes suggested their elaborations not only became more coherent by ensuring a tight alignment between main claims and a thesis statement but also increased the possibility of leading to a reasonable conclusion with a comparatively high level of logical strength. Meanwhile, it would be reader-friendly to understand relationships between ideas if sub-claims precisely supported the main claims and the chains of the main claims accurately supported the rational conclusion (Figure 10.1). This finding was in line with studies by Al-Shaer (2014) and Liu (2011), highlighting the effectiveness for EFL learners is to organise loose ideas into structured claims and distinguish between primary and secondary claims. As a result, it promotes learners' understanding of the meaning and potential associations (Rosen and Tager, 2014).

Figure 10.1 A model thinking map created by a student



Inference refers to searching the evidence to support arguments and drawing conclusions or determining opinions using appropriate reasoning (Figure 2.1). Two (out of 17, 12%) students expressed similar perspectives: thinking maps made them

think about examples to support their ideas (IS03, see below) and to describe issues following a reasoning sequence (IS16, see below).

> (IS03): " I think it (thinking maps) was practical and clear."
> Interviewer: "What do you mean by saying practically?"
> (IS03): "I mean, it fits in with what I am going to write."
> Interviewer: "So can I say you worked together and discussed what specific examples you can use in your essays?
> (IS03): "Yes."
> Interviewer: "Which aspect do you think becomes clearer?"
> (IS16): "What I mean is I know from which aspect I can write orderly when describing an issue."

The above quotes indicated that students noticed the power to work in groups to discuss and choose appropriate examples to support their argument and began to notice the importance of reasoning sequence. These elements were regarded as a premise to draw a logical conclusion. This finding echoed several studies stating thinking maps strengthen students' capacity or consciousness to organise ideas into coherent paragraphs and keep to the topic (Kaeppel, 2021; Lee, 2013; Liu, 2011; Pishghadam and Ghanizadeh, 2006). Therefore, the composition may become more systematical, organised and unified (Al-Shaer, 2014; Rahman and Ambreen, 2018). Although only a few students mentioned such information in this study, at least this was a beginning to consider related components leading to a reasonable conclusion.

In addition to the skills enhancement, thinking maps accelerated students' writing process (IS04 and IS05, see below) because they knew what and how to write. The following quotes indicated that thinking maps enhanced efficiency and effectiveness when completing writing projects (Liu, 2011) due to time-saving and decreased writing anxiety.

> (IS04): "I wrote quicker than before and not stuck while writing because I clearly knew what I was going to write."
> (IS05): "I think it helped. I just wrote what came into my mind in high school. I think I would write at a faster pace if I had a framework to follow. "

In this section, students believed that thinking maps improved their writing structure and organisation, serving as two influential indices for better writing performance. A

complete structure could be regarded as a premise for a better organisation. Namely, they needed to identify what pieces of content were essential for their compositions and then determine in what order they should present them (Cali, 2021). Both progress revealed the improvement of students' core skills of Interpretation, Analysis and Inference. On the one hand, they interpreted the involved information and identified the elements required to form inferential relationships to draw a reasonable conclusion. On the other hand, the enhanced skills answered CT questions about what to write, how to arrange the contents and keep consistent with the topic.

As discussed in section 8.5.1, students believed it was easier to write following a writing direction in high school. However, some of them reflected that such given directions restricted their thinking from organising their essays. They probably encountered a dilemma between a dearth of thinking practice in high school and CT requirements in the university when writing essays. Therefore, some expressed the greatness of thinking maps because this strategy provided opportunities to practice their thinking (IS11, see below).

> (IS11): "... I think when I was in high school, the structure was limited to three paragraphs because the writing directions offered three points. So the structure was fixed. ... While in university, we had to think all things because we were only provided with a title without any writing direction. ... I think it was good to practice now."

The quote indicated that these new univerisity students who were accustomed to writing compositions following the specified outlines in high school might find it challenging to design a structure and organise contents themselves in the current situation. Furthermore, their sole pursuit of advanced language in writing with rare consideration about exploring underlying associations between sentences or paragraphs may hinder their thinking development. However, their writing without thinking requirement in high school was not in accord with the teaching objectives of thinking cultivation in the current stage. In this light, the thinking maps strategy was regarded as a practical instrument for these new students in university to practice thinking skills in writing. By revising thinking maps collaboratively, they helped each other and were motivated to use CT to think more comprehensively, eventually resulting in in-depth learning (Papanikolaou and Gouli, 2013). Therefore, creating

thinking maps collaboratively to facilitate writing could be a transition stage for their individual maps' drawing.

10.3.2 Students' concerns when using thinking maps

Although twelve (out of 17, 71%) students explicitly claimed they had never heard of or created any thinking map before, only three (out of 17, 18%) acknowledged it was the most challenging strategy among three CT strategies in this study. It was still worth noticing their opinions taking account of potential space to be improved for future teaching's sake. It seemed that the quality was their primary concern. One student (IS09) worried about the map's contents: whether the standpoints were biased or consistent with the key idea. Another student (IS18) thought the teammates could not understand it because of poor idea visualisation. The third student (IS16) described confusion about applying it to writing execution, although his/her mind was supposed to become clearer than before.

> (IS09): "I was always worried whether my thoughts were biased or missed the key idea. ... Although different people had (a) different understanding about the specific topic, I still hoped there was a standard answer. For example, if your (the instructor) thinking map was different from mine, I thought I was wrong, even though I believed that mine was reasonable."
>
> (IS18): "I can understand my thinking map, but I don't think other teammates understood. For example, I just wrote a simple word 'school' as a branch, and I would add something as sub-branches to the branch of the 'school'. But I felt my teammates couldn't understand because I had just written a simple word of school."
>
> (IS16): "I was still confused sometimes when I wrote in detail, although my mind became clearer. I mean, I knew the outline and direction, but I still have difficulties when writing the specific contents."

Although only three interviewees mentioned their concerns about the thinking maps' quality, this comparatively lower number did not indicate it was easy to show the logical relationships among components of the argument. Several studies confirmed similar concerns about students' difficulties, such as a long period when drawing thinking maps (Bilik *et al*., 2020; Chen, S.-L *et al*., 2011), especially for the first time (Rosen and Tager, 2014). In fact, I found a considerable proportion of essays still had

poor organisation or structure because of poor thinking maps during my coding processes, such as biased claims, irrelevant evidence or poor coherence with the topic. Therefore, it was assumed that some students in this study were unaware of their deficiencies in thinking maps. Wang and Seepho (2017) attribute Chinese EFL students' challenges in creating thinking maps to a gap between high demand for using a creative ability and an absence of such a capacity because of the rote learning shaped by Confucian culture.

Students' concerns and thinking defects in their essays reminded me of providing scaffoldings for further effective learning, such as more detailed instructions to draw thinking maps and acceptable evaluation criteria as follow-ups. This reflection was in alignment with the studies of Chen *et al*. (2011), Daley and Torre (2010) and Lee (2013), articulating that writing instructors ought to provide explicit guidance and feedback to help students benefit more from thinking maps.

10.4 How Group Discussion Developed Students' Thinking

This section firstly reported the benefits of group discussion and what corresponding core thinking skills were inferred from such advantages. Then, it lists students' several dissatisfied points when participating in group discussions.

10.4.1 The core thinking skills development by using group discussion

Interviewees benefited from group discussions on the development of two skills: Evaluation and Explanation. The Evaluation refers to assessing the relevance to the thesis or the strength of arguments (Figure 2.1). At first, two (out of 17, 12%) students stated that it would be easier to detect errors when exchanging ideas (IS12).

> (IS12): " I think the nature of discussion in groups is to help others and at the same time to help myself. What I discussed was not always correct, and I helped myself when other people pointed out my mistakes."

This quote indicated that error correction through group discussion was bidirectional. Students could discern other peers' pros and cons. At the same time, their comments were evaluated by other peers. Their two-way comments exchange is an effective interactive correcting process. As a result, their evaluation skill was developed by mutually and spontaneously assessing the reliability of diverse perspectives.

On the other hand, more students (n=10 out of 17, 59%) stated group discussion allowed them to brainstorm all possible ideas (IS03, see below). This finding was in accord with the perspective of Al-Shaer (2014) and Zhong *et al*. (2021). They highlight the effectiveness of this CT strategy, enhancing learners' ability to brainstorm and generate more ideas. Specifically, students "learned about different opinions from other students" (IS05), were inspired "to think from diverse angles" (IS08), "explored new ideas that they never imagined" (IS14) and "exchanged their perspectives" (IS12). In addition, they tended to "feel less pressured to speak up about their thoughts without a teacher-led situation" (IS11) and thus was easier to "open their minds to extend more aspects" (IS07). This finding is confirmed by Wang and Seepho's (2017) research, reporting that students are more likely to exchange ideas freely if no teachers are standing by. This result means that student-led group discussion creates a relaxing atmosphere by freeing the anxiety of being judged by teachers. And this comparatively relaxed environment may benefit students' communication and participation (Poole, 2000).

> (IS03): "I think there was one word to describe the greatest benefit of it (group discussion), which was a brainstorming. That is we could create more fantastic ideas by talking with other people. "

The above quote revealed that students were inclined to think broadly and accurately on writing topics, especially in relaxed circumstances. One possible explanation was that they had to evaluate the credibility of the keywords and logical strength in branches of the thinking maps and provide their understanding or analysis when being required to make suggestions (IS11).

> (IS11): "It happened once. When I came up with my disagreement about the structure, the other teammates didn't realise that problem, and they seemed confused, especially the writer who took charge of the writing project. I tried to convince him by giving my reason why I thought the structure was defective and explaining my points of view again and again. Gradually, they understood what I meant and agreed with me at last. And the writer also revised the structure of his essay. It cost us so long time, but I felt so proud of myself to convince them."

Another possible explanation was they needed to explain how they understood the concepts or provide reasons to support their opinions when sharing or debating their different views of point. Although what they expressed was not always correct or

proper, at least they had one more chance to reflect on their reasoning by participating in the discussion (IS12). Their evaluation or reflection of other peers' perspectives may stimulate them to generate more ideas.

> (IS12): "I would like to point out some information that other people did not mention. The nature of the discussion was helping others while helping yourself because sometimes what I expressed may not be correct. But if I wanted to help other people and chose to speak out my opinions, and they could also point out the drawbacks of my opinions. In this way, they helped me too."

These two causes revealed that students needed to assess the reliability or validity of claims, visualise argumentative processes by thinking maps and express their attitudes, especially when they had a divergence. A similar view was shared by Kaeppel (2021) and Lee (2013). They claim that collaborative concept mapping reminds and stimulates students to critically evaluate other learners' maps, effectively engaging them in discussing the concepts or the inferential relationships between branches in the maps. Students, especially those who took charge of writing projects, cannot bypass a procedure of evaluating and selecting relevant information when facing diverse suggestions. After all, it was unrealistic to put all the information they discussed in writing. This ability to assess claims and arguments indicated students improved their Evaluation skills.

Brainstorming also improved the Explanation skill, which referred to stating reasoning results accurately, justifying reasoning procedures of evaluation and presenting reasons for supporting arguments (Chapter 2, section 2.2.4, Table 2.1). When brainstorming ideas, students had to justify their standpoints by showing their reasoning if they held different opinions and intended to convince other students to revise thinking maps (IS11). This capacity to display the justifying procedures or arguments revealed that students improved their Explanation skills.

> (IS11): " I thought his essay's structure was problematic. He disagreed with me, so I asked the other teammates' opinions, ... I think they did not realise what's wrong with it. ... I tried to convince them by paraphrasing my ideas. They gradually understood what I was trying to express, ... At last, that writer revised his structure."

Another manifestation was students' essays became more comprehensive (n=8 out of 17, 47%), attributing to group collective ideas instead of individual opinions (IS14, see below). This finding was in line with a result of Al-Shaer's (2014) research, proving a positive influence of group discussion on students by providing "richer ideas and thinking" (p.18). Through discussion, students "reduced the extent of biased perspectives" (IS12) and "complemented or deepened perceptions from different standpoints" (IS17).

> (IS14): "... I thought my essay became more comprehensive because of diverse standpoints to interpret the issues. Some of them were brand new, which never came into my mind before."

The above quote indicated that students were more willing to or had an increasing consciousness of assessing the reliability, rationality, and connections. They shared more novel ideas and more profound interpretations from instant mutual feedback when discussing together. This interaction could partly explain why students felt their writing improved with less biased, objective, and more comprehensive ideas (IS04, see below).

> (IS04): "(Through a group discussion,) I found my perspectives were partial. But I believed my contents would be more comprehensive if five of us thought together. And my essay would be better if what they thought were better than me."

In this section, students believed that group discussion makes their writing more likely to become comprehensive by brainstorming possible ideas. As project writers, they had to reflect on the details of thinking maps and their first draft by assessing the credibility of peers' suggestions and determining which recommendations they should accept to revise the maps. In other words, group discussion helped them evaluate and discern weak or irrelevant information (Kaeppel, 2021; Pishghadam and Ghanizadeh, 2006) when revising the thinking maps in order to illustrate a more logical relation among components of the argument or correcting the linguistic and thinking errors of the first drafts. While other students who took part in group discussions to revise thinking maps tended to evaluate the branches' reliability and logical relationships, express their attitudes, and offer modifications by explaining their reasoning. As Ikuenobe (2002) discussed, group discussion generates critical thinking by explaining,

evaluating and synthesising diverse perspectives. In this light, participants developed Evaluation and Explanation sub-skills.

10.4.2 Students' concerns when taking part in group discussion

Although ten students held positive attitudes towards the effectiveness of group discussions, nine (out of 17, 53%) displayed their worries or dissatisfaction, with the largest number of students having "negative" feedback among the three CT strategies. Firstly, five (out of 9, 56%) students thought their group members didn't participate actively (IS07, see below). For instance, some of them "did not treat it seriously" (IS02), "not prepare before the class" (IS04 and IS15), or "just sat there and didn't want to participate" (IS05). All their behaviours indicated that these students hold negative attitudes towards group discussion.

> (IS07): "Three boys talked more than the other two girls. They basically didn't say anything, and only three of us kept talking."

Secondly, three (out of 9, 33%) mentioned they didn't take full advantage of the discussion time (IS17, see below) because they thought "it was too long to talk about other topics" (IS06) or "the effective discussion time was short" (IS15).

> (IS17): "I thought we waste so long time in group discussion. ... I didn't think discussion time was so long for us. I mean, we didn't make the best of it. Sometimes we indeed needed so long time to discuss, but sometimes... It's hard to say. ... Some groups took advantage of the discussion time on debating the contents of the thinking map. However, some groups just ended the discussion because they failed to recognise some drawbacks of the map."

Thirdly, four (out of 9, 44%) expressed the dissatisfaction about their harvest from discussions: they "couldn't get more inspiration from each other" because they were at "a similar level" (IS09); therefore, their conversations were "at a superficial level" (IS04). Or they were "not accustomed to this learning format" because "they never experienced group discussion" (IS02). Or it seemed that "what we concerned about were not the same, like comparing apples to oranges" (IS09). One student (IS13) questioned the effectiveness of discussing in groups when revising thinking maps. "Students who created their thinking maps had their thoughts, so it carried little weight to help determine the thinking line for writing", which could be a possible

justification to explain why some students thought "the discussion time was too long" (IS06).

In a sentence, students' dissatisfaction with group discussion concerned inactive participation: the less effective discussion time and the limited achievements. All these drawbacks were relevant to the effectiveness of group discussions which was a universal challenge in teaching practice (Dallimore *et al*., 2004). In addition, these unsatisfied experiences may negatively influence the final learning outcomes. Khodadady and Ghanizadeh (2011) state a similar point of view: learners' discussing other topics or uncooperative attitudes in discussion failed to achieve the effectiveness of group discussion. Although students had unsatisfactory points in group discussions, there was no denying that group discussion effectively improved their CT skills. At the same time, instructors should pay great attention to students' feedback for better teaching in the future.

10.5 How Peer Review Developed Students' Thinking

As reported in the previous sections, this section follows the same structure of presenting the pros of peer review and the corresponding developed core thinking skills. Then, it lists the dissatisfactions students are concerned about when conducting peer review.

10.5.1 The core thinking skills development by using peer review

In the peer review strategy, all participants played a dual role of being writers to receive feedback and reviewers to provide comments when revising. Students in this study developed three core skills: Analysis, Evaluation and Self-regulation. In the Analysis and Evaluation skills field, twelve (out of 17, 71%) students said peer review helped them detect the drawbacks they couldn't recognise in their compositions (IS04, see below).

> (IS04): "It (peer review) absolutely worked. I think It's hard for me to find errors in my writing, even though I read it for several times. But they (the teammates) can find them (the errors) at first sight."

This quote presented it was easier for reviewers to find errors or mistakes from the third party's perspective than for the project writers who took charge of writing

projects. In addition, it reminded students to focus directly on the inappropriate places they failed to find in the compositions, which saved their time and energy for better writing.

Being reviewers, they tried to detect flaws in two strands: language use and critical thinking. In terms of linguistic errors, all interviewees (n=17 out of 17, 100%) reported that they typically corrected grammar points and suggested advanced vocabulary and sentence patterns. Regarding critical thinking skills, twelve (out of 17, 71%) students said they revised or corrected thinking mistakes. Six (out of 12, 50%) said that they revised coherence (IS11, see below), four (out of 12, 33%) modified linking devices (IS14, see below), and three (out of 12, 25%) changed logical issues in writing (IS18, see below).

> (IS11): "...I preferred correcting their essays in structure rather than some details. For example, I liked to point out the incoherent places among paragraphs."
> (IS14): "... The writer used the linking word *so* or *therefore*, but he did not write the causal relationship. Or there was a gap between two paragraphs, and the writer should add one more paragraph. Or the thinking was incoherent."
> (IS18): "I helped them find the logical problems in writing."

The quotes indicated that reviewers analysed and evaluated whether project writers used the language correctly. Meanwhile, they identified and assessed the claims or arguments in the construction and organisation. In this sense, viewers developed their Analysis and Evaluation skills.

In terms of the Self-regulation skill, it means self-examination of reflecting on viewers' suggestions as well as discovering problems and self-correction of the errors they made based on their reflection (Chapter 2, section 2.2.4). On the one hand, twelve (out of 17, 71%) project writers considered the received feedback and determined whether they accepted the flaws found by the reviewers (IS12 and IS17, see below).

> (IS12): "... what I read and they (teammates) read about my essay was quite different because our thinking was diverse. For example, I thought the sentences were logical and connected. However, they could find a gap among the sentences and failed to understand what I wanted to express. I had to reconsider whether their suggestions were right or wrong. "

> (IS17): "... Definitely it (peer review) helped if someone else revised for you. I quite understand the meaning of the sentences in my writing, but some of them couldn't. If most of them couldn't understand, I thought it was my problem, and I needed to rethink. So I thought my composition was read more objectively by peer review."

These quotes displayed that peer feedback enabled students to analyse the essay generally or specifically (Silva *et al*., 2016). Meanwhile, student writers reflected on their reasoning process again and checked the errors or mistakes in language. Furthermore, they integrated constructive feedback and eventually determined whether they applied it in the second draft to improve writing quality. This reflection echoed the studies of Alharbi (2018), Kölbel and Jentges (2018) and Silva (2016), claiming students who receive peer feedback learn to rethink the shortcomings of the essays and correct the errors or mistakes. In this light, students developed a Self-regulation skill by self-examining the feedback and self-correcting the compositions' drawbacks in this reflective process.

On the other hand, seven (out of 17, 41%) students stated that peer review provided more opportunities to learn from others (IS01, see below).

> (IS01): "Well, I benefited from reading other people's writing because I can find some good points in their compositions."

Firstly, three out of seven mentioned that it improved viewers' linguistic abilities, especially in vocabulary and grammar, because they have more opportunities to read other people's essays (IS04, see below).

> (IS04): "I can learn some grammar points by reading through other students' writing, especially those I was not familiar with. I took notes when reading some advanced grammar points or sentence patterns.
>
> (Interviewer): 'What else?"
>
> (IS04): 'Some sentence patterns. or some long or unfamiliar words.'

Secondly, the other three students thought it reminded them of avoiding making the same mistakes as the writer (IS17, see below).

> (IS17): "I would check whether there was any problem of the contents and I also reflected on whether I made the same mistakes. For example, some students would like to use some advanced vocabulary for advanced vocabulary's sake and the sentences were not reader-friendly and lack logic. I think I never did it in my writing."

The above three quotes indicated that students increased their consciousness to think reflectively. Namely, viewers could apply what they learned from revising student writers' essays to their own writing, such as advanced vocabulary and sentence patterns. Additionally, they learned a lesson by detecting some drawbacks and avoiding making the same mistakes. Therefore, peer review was regarded as a bidirectional channel to rethink and revise the writing when taking turns as the writers and reviewers. In this way, students improved the Self-regulation skill by self-examining whether they made the same mistakes as the writers. Once they made it, they can adjust the deficiencies in time for higher writing quality. Articulating a similar view, Beiki *et al*. (2020) point out that reflection and self-reflection were vigorous processes that could stimulate students to think critically about what they did before.

In a nutshell, students developed their Analysis, Evaluation and Self-regulation skills as both a reviewer and a writer through peer review. On the one hand, they analysed and evaluated the deficiencies in the essays, such as language, structure or organisation, to give the feedback. Additionally, they learned good vocabulary or sentence patterns when reading other students' compositions. On the other hand, they took into account their group members' feedback and corrected the errors and mistakes in their writing.

10.5.2 Students' concerns when taking part in peer review

In this study, six (out of 17, 35%) students reported three factors impeding their writing development. At first, four students stated that their capacity was limited in recognising or correcting errors (IS01 and IS06, see below), especially linguistic mistakes or using the checklist to review their writing in thinking.

> (IS01): " Well, my English level did not allow me to find their grammar mistakes. Even I did, sometimes I didn't know how to correct it. So I felt it was hard for me to review their essays."
>
> (IS06): " If you wanted to correct the wrong sentence, you must have a higher ability to find it and provide a better one. Sometimes, I cannot make it."

These quotes indicated that some learners' English level, esp. in language use, restricted them from detecting or correcting drawbacks. This phenomenon was quite

common for foreign language learners, although their English level was at least above medium level (section 9.2). It can be imagined that this difficulty would be more severe for EFL students with lower English levels. In addition, students' English language proficiency indeed affected the quality of feedback (Han and Hyland, 2015; Oksana and Deoksoon, 2017). Therefore, they naturally expected to acquire more linguistic knowledge. Students in this study reviewed peers' compositions individually after class, and they had another opportunity to share their comments or discuss their uncertainty with group members in class. Their oral exchange could partly help figure out their difficulties or reduce misunderstandings when reviewing the papers. This interactive feedback between the feedback providers and between feedback providers and receivers was more likely to increase the feedback quality than a one-way information transmission from the providers to the receivers (Yu *et al*., 2021).

Another point student reported was that someone in his/her group held a negative attitude and was reluctant to review the essay (IS02, see below). The following statement indicated that some students had inactive, even resistant attitudes towards peer review or were less motivated to review the essays. In addition, they may lack a cooperative spirit contributing to peers' writing improvement. Meijer (2020, p.1227) interprets this behaviour as "free-riding and sucker-effect" in this collaborative and communicative peer review, no matter whether this inactive student does it intentionally or not. The "casual" attitude towards the feedback process may reduce students' writing motivations because they cannot receive adequate feedback as writers and may refuse to provide their comments as reviewers due to dissatisfaction with the feedback providers (Yu *et al*., 2021). Thus, students who contribute more to completing projects even feel frustrated owing to their less cooperative peers (Zhai, 2021). Tsui (2002) further elaborates that they positively influence building peers' confidence only when students are cooperative and supportive.

> (IS02): "... Well. Some students were very responsible, and they pointed out many mistakes and provided their amendment. While some of them said they forgot to review and quit our discussion. Maybe they really forgot. ... I think they treated it (peer review) as a burden instead of a way to improve English learning, so they did not want to take part in. ... How to say? Actually, once we decided who the writer was, some of them actually didn't care how to finish that job. They thought it was their (the writers') own business. "

> (Interviewer): "... I thought I should help you realise the importance of paper reviewing and comments sharing and the other important thing was that you were a team."
>
> (IS02): "That is a cooperative relationship."

Finally, one student reported that their feedback may hurt other students' self-esteem because they were a student-student relationship, not a teacher-student (IS11, see below). This quote indicated some students' hesitation when peer reviewing the compositions, especially when intending to provide negative comments. The feedback providers were worried their negative suggestions may influence their impression of the feedback receivers. Face-saving was essential for some Chinese, especially for some excellent Chinese students. They always wanted to leave a 'perfect' impression on their peers and teachers. Therefore, it was necessary to consider the potential risk of hurting both providers' and receivers' emotions and reducing motivation to write revisions. This finding seems to be consistent with other studies which claim negative emotions (Yu *et al*., 2021), such as an embarrassment for writers (Vorobel and Kim, 2017; Zhang and Hyland, 2018), sensitiveness out of intolerance towards the feedback (Alharbi, 2018) and discourage students' engagement. Once students had this tendency, teachers could encourage them to reverse these negative feelings into incentives, and some students indeed achieved it (Han and Hyland, 2015). Students who suffer from negative emotions are suggested to learn to analyse and evaluate the feedback from diverse sources to improve their writing (Yu *et al*., 2021).

> (IS11): "We were students, not a teacher. It's OK if their expressions were not wrong. Otherwise, It looked like a show-off if I revised their middle level of words or sentences to advanced ones. It's not appropriate giving them a sense of superiority because we were equal."

In this section, students reported their three concerns when doing peer review: the inability to revise, the uncooperative attitudes and their worries about raising negative emotions for feedback receivers. These negative feedback and experiences were consistent with the study by Yu *et al*. (2021), claiming that Chinese EFL learners who have negative experiences employ low motivation or inactive engagement. The "negative feedback shaped the relational landscape in which L2 writers interpreted and used feedback" may hinder them from using this strategy (p.3). Therefore, teachers and students should be alert to such adverse effects in future teaching.

10.6 Summary

This chapter provided findings in tandem with the third research question of how three CT strategies impacted students' CT development. Broadly, thinking maps, group discussion and peer review developed students' thinking from different perspectives, including all the core thinking skills of Interpretation, Analysis, Inference, Evaluation, Explanation, and Self-regulation. This result validated and supplemented generic results of CCTST in Inference, Analysis and Evaluation. Specifically, Analysis and Evaluation were the comparatively most involved skills based on the results of four sources of the CCTST, students' writing, questionnaire and interview.

Many scholars prove these CT strategies' effectiveness in improving critical thinking (Al-Shaer, 2014; Beiki *et al*., 2020; Dallimore *et al*., 2004; Darani and Nafiseh, 2019; Kaeppel, 2021; Rosen and Tager; 2014, Tsui, 2002), especially when the participation is in a collaborative way (Asrita and Nurhilza, 2018; Mujtaba, 2021). One plausible justification is that working in groups triggers mutual scaffolding (Li and Zhang, 2021) to communicate and learn from each other in a supportive environment (Darani and Hosseinpour, 2019).

Although students expressed different concerns or challenges when engaging in these strategies, the most prominent issue in this study was the low efficiency of their cooperation, especially in peer review and group discussion, which was a common challenge for collaborative learning. These results will inform my future teaching.

Chapter 11 Discussion

11.1 Introduction

This chapter aims to discuss the previous three chapters' findings on the applicability and effectiveness of CT-oriented PBL classes to improve students' CT. Chapter 8 accounted for students' CT development after PBL intervention. It is evident that students improved generic CT skills in the CCTST and content-specific thinking skills in writing. Whilst the interview findings of their changed writing standards after PBL intervention indicated that they generally enhanced thinking dispositions. In addition, they developed four out of six core thinking attributes in the CCTDI, such as Analyticity and Confidence in reasoning. However, their development was unbalanced with high Inquisitiveness and low Truth-seeking and Open-mindedness. Chapter 9 reported that students held a positive attitude towards the effectiveness of CT strategies and PBL intervention to improve CT, indicating they believed CT-oriented PBL pedagogy enhanced their thinking. Moreover, the strategy of thinking maps was the most effective. Chapter 10 explored how students developed their CT by presenting the benefits of each strategy and highlighting two primary concerns when using these strategies including their limited competencies in English proficiency and CT and the low efficient group work. In brief, the above findings demonstrated the applicability and effectiveness of CT strategies and PBL teaching intervention in Chinese EFL classes, indicating the possibility of teaching CT in Asian EFL classrooms across tertiary education.

In this discussion chapter, the research questions are directly answered by discussing the findings of the previous three chapters and linking the findings to previous studies. Firstly, two learning contexts focusing on improving CT and examination performance are described in order to portray how different contexts positively or negatively influenced students' CT learning. Then, the discussion moves to present how CT-oriented PBL teaching intervention succeeded in enhancing students' CT learning outcomes. Finally, there will be an examination of students' CT performance, attitudes and perceptions.

11.2 Context for CT Development

It is possible to teach higher-order thinking skills in a foreign language class (Bağ and Gürsoy, 2021; Ding, 2016; Soufi and See, 2019) in the Asian EFL teaching context. The findings in this study confirmed the feasibility of teaching CT explicitly in EFL classes. This section presented how a PBL course design embedded with CT instruction created a CT-oriented context in the CE course and how CT strategies and different learning contexts impacted students' thinking outcomes. In addition, an assessment-oriented context of encouraging students to obtain high scores in exams which hindered students' CT learning is illustrated.

11.2.1 A CT-oriented PBL course design

Participants in this study held a positive attitude towards the effectiveness of PBL teaching intervention (section 9.3), indicating they believed this pedagogy benefits their thinking improvement. This result supports Handhika *et al*. (2018), who suggest the project design should allow students to think deeply and reflectively, model how to think critically and provide feedback on group discussions. These benefits are barely achieved by traditional teacher-centred pedagogy, which was reported to hinder Chinese students' thinking enhancement (Praba *et al*., 2017; Tang and Biggs, 1996; Zhao, 2020). The current study extends the project design to a course design involving a teaching objective and CT instruction.

11.2.1.1 An explicit CT teaching objective

Although PBL employs a high probability of thinking enhancement, it does not guarantee it always works (Handhika *et al*., 2018; Zare, 2015). Previously literature (Abrami et al., 2008; Sun, 2017; Tsui, 2002) suggested the most effective way of improving students' thinking skills is to teach CT explicitly by setting up it as a learning outcome in EFL class. Therefore, it should be involved in the PBL course design. Specific to this study, besides language development, CT was separately listed as one teaching objective due to the demand for thinking training, allowing teachers to have a clear direction. It is suggested that "making CT requirements a clear and important part of course design is associated with larger instructional effects" (Abrami *et al.*, 2008, p.1121). In this sense, a clear CT teaching objective is an essential starting point for teaching thinking.

11.2.1.2 CT-enhanced instruction

After setting a clear teaching objective, what CT instruction and how to use it in teaching need to be considered next. Otherwise, it achieves negligible results with only an intention to list CT development among the course objectives (Abrami *et al.*, 2015; Niu *et al*., 2013). As discussed in Chapter 4, section 4.2, CT instruction included lectures and CT strategies. In terms of the lecture, teachers determined what content knowledge should teach situated in the CE course. Pally (2001) suggested that it contains linguistic and subject-based CT knowledge. Although Pally's work was published 21 years ago, his suggestion still applies to the current study. Specifically, the linguistic knowledge contained vocabulary, grammar and language skills; CT knowledge involved thinking skills in writing and how to use CT strategies or auxiliary instruments (i.e. a rubric of assessing thinking skills in writing). Firstly, the current study attached importance to content-specific thinking skills in EFL writing in the Chinese context. In contemporary literature, there is no denying that some common CT skills and dispositions share in various subjects (Moore, 2013). However, critical abilities are subject-specific (Ennis, 2018), and the same CT elements hold different places in different disciplines (Cáceres *et al*., 2020; Moore, 2013). Therefore, it is assumed that the content of specific thinking skills in language abilities, such as reading or writing, is different. Only if students are taught particular items for each core thinking skill can they use them to write and comment. In addition, teachers could achieve thinking teaching objectives for these first-grade students if the items are designed based on the writing requirements of the CSE 5. Therefore, teachers must define the core thinking skills with related items specific to the EFL writing requirement of CSE 5. In this study, CT skills were categorised into six core thinking skills: Interpretation, Analysis, Evaluation, Inference, Explanation, and Self-regulation, with fifteen descriptions (Chapter 2, section 2.2.4, Table 2.1). Since the CCTST (the simplified Chinese version) only assesses Analysis, Evaluation and Inference, this study simplified the above fifteen descriptions into ten under the targeted three skills in order to compare students thinking skills' performance in the test (CCTST) and writing (Appendix I).

Take unit 3 as an example. Students learnt expressions about different life stages of older and younger generations in western countries and a sub-skill of comparison and

contrast. When thinking about the writing title, students who took charge of writing extended that topic by comparing or/and contrasting two generations between their parents and themselves in a Chinese context. When drawing thinking maps, they used keywords to describe differences or/and similarities, analyse why they held different attitudes towards the same issues, and conclude and reflect on what they should do as a younger generation when having divergent perspectives from their parents. Some of them also further expressed their expectations of their parents. As a result, they generate a new and deep understanding of the generation gap by utilising the thinking skill of comparison and contrast.

In terms of CT strategies, thinking maps, group discussion and peer review were used in this study. When teaching CT strategies, I modelled how to draw a thinking map (Appendix L1), provided a thinking checklist to help detect thinking errors for peer review (Appendix I) and allotted time for students to discuss the deficiencies in thinking maps and mistakes of language and thinking in class. Student thinking practice in different stages of PBL (i.e. preparation and implementation stage) raised their thinking awareness and more likely to apply thinking skills to writing. For example, using the thinking map to prepare for writing enabled them to have a clear mind before writing and thus, increased their learning efficiency. Group discussion brainstormed their ideas, allowing them to think from diverse perspectives when implementing projects. This thinking process made them more aware of thinking. As a result, they added CT as a new criterion when evaluating writing quality (Chapter 8, section 8.5.2). In addition, the higher mean scores in the CT test and increased frequencies of using thinking skills in writing revealed that CT instruction contributed to thinking development. These positive teaching outcomes are in line with SCT, namely, mediation (i.e. CT instruction) increased learners' cognitive competence through knowledge acquisition and internalisation (Chapter 4, section 4.2). These results also confirm a meta-analysis study by Abrami *et al*. (2008), advocating that CT teaching principles should be explicit if practitioners integrate thinking skills into subjects. This explicitness resides in a CT-enhanced course design or integrating relevant teaching components (Niu *et al*., 2013) to maximise CT teaching outcomes (Abrami *et al*., 2008). Moreover, the positive result supports the perspective of the possibility and effectiveness of integrating CT into an EFL subject in the Chinese

context (Lin, 2014; Sun, 2017; Zhang *et al*., 2020). Cáceres *et al*. (2020) evaluated the integration way as highly cost-effective in fostering students' thinking, especially for teaching programs without a separate CT course.

Finally, CT-enhanced instruction activated students' active learning. Although showing some passiveness in the learning process, students presented their initiative in two facets. At first, they are free to choose any title they are interested in. By means of actively engaging in learning out of interest, students are more likely to employ higher-order thinking skills to explore deep knowledge with higher motivations (Kwan and Wong, 2015; Tsui, 2002). Their interest in topics around their lives may stimulate them to critically reflect on their authentic problems, resulting in either a potential promising thinking performance in writing or a reasonable option for their problem-solution. Their learning became more meaningful, being the start of active learning given the situation of the participants in this study, who were always passively given titles and outlines to write in high school. To select a proper title to write, they not only had to consider their interest or concerns related to the topics they learnt in class but also think about the possibility of extending the contents. In addition, they had to work in groups to determine what they should discuss in the group discussion when revising the thinking maps and essays (this is discussed further in section 11.2.2). As a result, teachers transformed their roles from a leader to an assistant to support learners to meet the demands of project completion. Compared to a traditional teacher-dominated pedagogy without enough freedom to make decisions (Li *et al*., 2016), PBL teaching allows students to be freed to choose the writing topics and determine what they need to discuss. This freedom benefited their independent learning abilities and provides an active learning environment.

This section discussed a CT-oriented PBL course design involving an explicit CT teaching objective and CT instruction. Students were proved to input enough knowledge prepared for completing the writing project. The following section will introduce how thinking strategies are catalysed to output this knowledge.

11.2.2 CT strategies as a catalyst in PBL contributing to CT development

Students held a positive attitude towards the effectiveness of CT strategies, indicating they believed thinking maps, group discussion, and peer review improved their CT.

This finding is consistent with several studies, claiming the above three CT strategies help students think critically in different aspects (Alharbi, 2018; Al-Shaer, 2014; Cali, 2021; Daley and Torre, 2010; Lee, 2013; Liu, 2011; Kaeppel, 2021; Khodadady and Ghanizadeh, 2011; Kölbel and Jentges, 2018; Nussbaum, 2002; Rahman and Ambreen, 2018; Silva *et al*., 2016). This study selected these strategies based on the demands of the writing project and thinking development.

In the CT-oriented PBL course design, an added CT teaching objective and the lectures are stepping stones to address critical thinking issues. However, merely using a mentoring approach (i.e. imparting CT knowledge) does not produce notably strong results unless combined with other methods (Abrami *et al*., 2015). It implies that lectures enable students to acquire knowledge; however students may still feel challenged when using thinking skills to complete projects and they need mediation, such as CT strategies, to help them apply the acquired knowledge to complete projects. In this light, it is necessary to integrate CT-enhanced strategies into different PBL stages (Figure 4.1) to offer students more opportunities to practise thinking and activate their mental processes in carrying out projects. As one major characteristic of PBL teaching, group work leads to better CT performance (Abrami *et al*., 2015; Cáceres, 2020; Niu *et al*., 2013; Wang, 2017), allowing students to help each other, especially when lacking CT and PBL learning experience. Therefore, building a collaborative context to input and output relevant knowledge is essential. This perspective is supported by a ZPD theory: working in groups could fill the knowledge gap between learners' inabilities and the assistance of teachers and "student experts" (Chapter 4, section 4.2).

Instructional language and CT strategies of thinking maps, peer review and group discussion were selected to present students with different ways to use thinking skills and increase their CT awareness. These strategies work together and play an integrative role in making up for the deficiency that any individual strategy fails to achieve. After all, a single method is usually unexpected to develop every core thinking skill. This section first discusses the feasibilities of using strategies to improve thinking in PBL and then investigates common challenges students meet when using these strategies.

11.2.2.1 Feasibilities of CT strategies

Given that most practitioners lack CT professional knowledge (Pithers and Soden, 2000; Snider, 2017; Zhang *et al*., 2020), it is more practical for them to achieve teaching outcomes if they alter and apply commonplace strategies to their teaching practice (Tsui, 2002). As universal instructional methods for the pre-writing and post-writing stages, thinking maps and peer review strategies were selected for students to prepare and revise writing individually. In addition, as a backup for the above two strategies, group discussion was employed to increase the quality of the thinking maps and the writing draft. These three strategies are also suggested by the CSE 5 as students' writing expression strategies of planning and revision (NEEA, 2018).

Firstly, this study found that the thinking maps strategy is feasible for teaching CT. Students acquired relevant knowledge through teachers' lectures and modelling to draw the maps (Appendix L1). Through presenting the structure and organisation of passages in the textbook, students learnt how those authors organise their writing from a reader's perspective. Then, they were required to draw thinking maps to outline the structure and organisation of their draft from a writer's perspective (Appendix L2). By drawing thinking maps, they developed the application skill.

In this study, students benefit from thinking maps by allowing them to have a clear mind about what to write and making the writing more organised and coherent, suggesting the development of Interpretation, Analysis and Inference skills (Chapter 10, section 10.3.1). Although fewer studies classify the advantages of thinking enhancement into specific core thinking skills, many scholars present similar perspectives by reporting its effectiveness and necessity as a pre-writing strategy (Al-Shaer, 2014; Lee, 2013; Liu, 2011; Rahman and Ambreen, 2018; Wang and Seepho, 2017). Therefore, this strategy could also be employed in teaching thinking as a universal writing teaching method. Additionally, students' most positive attitude towards its effectiveness (Chapter 9, section 9.2.2.3) indicated that they benefit the most in this study. The cause why this strategy helped them the most was they were always given a writing direction without considering the structure and organisation before writing in high school (section 8.5.1), so it was challenging for them to achieve it in university. Drawing thinking maps allows them to practice thinking abilities,

accelerates their writing speed and reduces their writing anxiety (Chapter 10, section 10.3.1). Without it, they may struggle with what to write or how to organise the essays. In this light, integrating thinking maps into PBL teaching is feasible from theory and practice.

Secondly, this study also found peer review strategy was essential. After completing the first draft, the student who drew a thinking map acting as a writer received comments, while peer reviewers provided their feedback (Appendix M). Therefore, every student employed equal opportunity to be a project writer to reflect on the comments and self-correct the errors and a peer reviewer to analyse and evaluate the essays (section 10.5.1). Most interviewees (71%) thought this strategy helped project writers detect the drawbacks they failed to recognise. At the same time, they benefited from self-examining and self-correcting the received suggestions. These advantages illustrated they developed Analysis, Evaluation and Self-regulation skills (Chapter 10, section 10.5.1). This finding aligns with several previous studies claiming that peer review helps students reflect on their shortcomings (Kölbel and Jentges, 2018; Ozogul and Sullivan, 2007), especially when they have to refine or rewrite their essays (Tsui, 2002). In addition, being reviewers, students have to employ specific skills, for instance, analysis and evaluation, to provide comments (Ozogul and Sullivan, 2007; Silva *et al*., 2016).

As a widely-used strategy in post-stage writing teaching, this study's CT-enhanced peer review process takes participants through two loops. The first is offering feedback after class individually, and the second is their collective communication by discussing and exchanging their suggestions which will be addressed in the next section. These two loops combination provided project writers and reviewers with the first opportunity to analyse and evaluate the essay by themselves and then the second opportunity to exchange and reflect on suggestions from other peers' comments. Therefore, it is interactive from multi-directions between reviewers and project writers and among the reviewers. Although some students met some challenges, such as insufficient linguistic and thinking abilities to detect or correct errors, or peers' negative attitudes (Chapter 10, section 10.5.2), it is regarded as the interviewees' favourite strategy in this study (Chapter 9, section 9.4). Meanwhile, their positive attitude towards its effectiveness (Chapter 9, section 9.2.2.3) indicated that they prefer

others to comment on their essays, contributing to thinking development. Peer review, another common language writing teaching strategy, enhanced students' thinking development by providing more reflective opportunities to exchange comments. In this sense, it could be applied to PBL classrooms aimed at thinking development.

Finally, as a collaborative method, group discussion was integrated into the preparation and implementation stage of PBL (Chapter 4, section 4.1, Figure 4.1), supporting the other two strategies. After finishing the individual work, it provided a second chance for students to collectively review, discuss and reflect on the quality of thinking maps and peer review results. Without group discussion, their thinking practice is one-sided lacking instant interaction. Therefore, group discussion allows project writers to reflect and be inspired to generate more ideas by communicating with reviewers instead of passively receiving the feedback. This also gives reviewers one chance to discuss the thinking maps and the first drafts with project writers and other reviewers. In this way, each student has multiple opportunities to practice and develop thinking as a writer and a reviewer, although they only write one essay project for the whole semester.

Since different groups had different problems or challenges, what students discussed was almost tailor-made and their learning became more active by detecting and solving problems. No matter what they discussed, the general topic was about how to make thinking maps better for more precise writing directions (i.e. the focus question and logical relations) and how to make the feedback more reliable and easier to revise (i.e. sharing feedback, exchanging and evaluating suggestions). In this sense, the strategy of group discussion in the present study is student-led, followed by a general direction of talking about their comments or confusion in thinking maps or revising drafts. Following thinking maps and peer review strategies, group discussion makes both strategies employ collaborative and scaffolding characteristics. This collaboration contributes to a better learning outcome than learning individually.

Students' positive attitudes (Chapter 9, section 9.2.2.3, Table 9.12) indicated that they believed group discussion helped them think critically. Brainstorming is a major benefit that interviewees mentioned (Chapter 10, section 10.4.1). They felt it generated more ideas by exchanging and learning different opinions from peers (IS05, IS12), discovered novel ideas that they had never thought of before (IS14) and

thought from diverse directions (IS08). These results are in agreement with studies (Brookfield and Preskill, 1999; Darani and Hosseinpour, 2019), claiming that it provides a supportive context for students to develop critical understanding, inspire more ideas and appreciate these diverse perspectives from different backgrounds or experiences. These benefits eventually lead to learners' CT development and better academic performance (Asrita and Nurhilza, 2018), especially when the teaching goal is developing higher-order cognitive skills (Tsui, 2002).

Being a complementary strategy to the other two methods, the combination of group discussion and thinking maps and the integration of group discussion and peer review work better than independently. Each combination creates a two-way information transmission to share peers' perspectives, receive instant feedback and discuss their confusions, misunderstandings or uncertainties. Their interactive dialogues involving more critical discussions trigger deep thinking about the discussed problems and facilitate problem-solving in revising thinking maps and sharing feedback. Once solving problems, they may perform better in their writing projects. Mujtaba *et al*. (2021) explicitly pointed out that collective revision outperforms the amount of settled and abandoned mistakes. Daley and Torre (2010) shared a similar view claiming that students do not perform better by simply drawing thinking maps without feedback. In contrast, one-way communication may restrict their access to feedback or revision (Yu *et al.*, 2021), resulting in a comparatively low quality of thinking maps and draft revision.

Storch and Wigglesworth (2018) stressed the necessity of introducing group work to brainstorm ideas before writing and gain feedback from teachers or peers on the writing. As a further step of thinking maps and peer review, group discussion existed throughout nearly all the stages of PBL, reminding participators of stressing CT all the time. In this sense, group discussion is indispensable for this CT-enhanced PBL course design.

This section has discussed the feasibility of integrating three strategies into CT-oriented PBL teaching design. Thinking maps and peer reviews help students plan to write systematically and receive constructive comments for better thinking performance in writing. Group discussion gives them more chances to evaluate and reflect on providing and receiving feedback, which contributes to revising their

thinking maps and more reliable comments and eventually motivates their actual thinking (Bağ and Gürsoy, 2021). In this light, the triple strategies improve their content-specific CT outcomes in writing. This finding is in agreement with those obtained by Morse and Jutras (2008, p.243), pointing out that "concept maps with feedback produced a measurable increase in student problem-solving performance and a decrease in failure rates". Therefore, teachers should encourage students to critically discuss their peers' thinking maps before taking part in an orderly writing activity (Lee, 2013). As a result, they complement their benefits and work together to improve students' performance.

Besides, the triple combination also triggers collaborative scaffolding among students to support each other (Li and Zhang, 2021) and come to grips with difficulties in writing. Although instructors impart language and CT knowledge in class, some students may still feel challenged when individually drawing thinking maps or reviewing essays after class due to their inexperience. These collaborative strategies enable learners to balance the gaps in their diverse learning abilities, linguistic proficiencies and thinking levels, eventually improving their collective cognitive critical thinking skills and dispositions. Therefore, CT strategies, as effective instruments to apply the CT knowledge, acted as a catalyst to expand PBL's function in thinking development. In addition, students took more responsibilities when using these three strategies, indicating that the centre of the classroom transferred from the teacher to these student participants.

11.2.2.2 Challenges of CT strategies

As collective scaffolding, CT strategies allow students to analyse, synthesise, and evaluate ideas collaboratively and stimulate active learning, resulting in thinking enhancement. However, some obstacles that may suppress their CT learning should not be neglected. Based on the results of students' concerns when using CT strategies (Chapter 10, section 10.3.2, 10.4.2, 10.5.2), they were twofold: their restricted capacities and low efficiencies of group work.

Firstly, students' deficiencies in language or thinking may restrict their performance of thinking maps or error corrections. When working individually, some participants may fail to present their thinking maps to peers because of poor visualisation or have

difficulties applying them to writing practice (Chapter 10, section 10.3.2). Even working in groups, they sometimes felt it difficult to detect subtle mistakes or offer more appropriate modifications owing to their similar English abilities. These deficiencies made their discussion "at a superficial level" (IS04). Yu *et al*. (2021) articulated a similar finding, pointing out that low-quality feedback is universally observed in peer evaluation, such as wrong or superficial feedback. Therefore, it is natural for some participants to believe they made limited progress (Chapter 10, section 10.4.2). Insufficient language proficiency is universally reported to hinder CT performance (Fong, 2003; Jin and Cortazzi, 2006; Kirby *et al*., 1996; Koda, 2005; Robertson, 2000; Zhang *et al*., 2020; Zhong and Cheng, 2021). Therefore, assuming that they could achieve a better CT correction outcome if their language abilities improve is feasible. Otherwise, their inadequate foreign language level may militate against their idea expression, even if they developed their thinking. In addition, low language proficiency impacts students' confidence in providing feedback (Cho, 2017; Zhai, 2021), explaining why some students in this study held a negative attitude towards peer review. Given the reality of their inability, EFL teachers' responsibility is to help students enhance their thinking skills while teaching them language (Shirkhani and Fahim, 2011).

Secondly, low efficiency in group work may negatively influence their group discussion and peer review performance. Some participants expressed disappointment with their partners' inactive participation by stating "they did not treat the discussion seriously" (IS02) or "they just sit there and say nothing" (IS05). As a result, their "effective discussion time was too short" (IS15). Their dissatisfaction indicated that these students might not be accustomed to studying collectively and need more time to learn how to cooperate. Or they did not take CT learning seriously.

In brief, their focus on one strand of language instead of language and thinking correction when revising essays (Chapter 8, section 8.5.3), passive attitudes towards peer review, and insufficient language or thinking abilities made them detect fewer errors than expected. For example, IS09 questioned when reading the feedback: "I am afraid they can point out or correct all the mistakes in my essay." Although more corrections do not signify a high quality of feedback, it may at least indicate whether they developed thinking skills, adapted to the new standard to discern more thinking

errors or increased their thinking awareness. In this light, spotting more mistakes may present these first-grade students could apply the newly acquired thinking knowledge to evaluate the essays. If failing, some groups may make the least allocated discussion time in class because they did not have much information to discuss. As a result, they "talked about other topics" (IS06), an unfulfilled effect of the pre-writing strategy of group discussion (Khohsima *et al*., 2016). In addition, this unfulfillment makes students who contribute more to the group discussion or with high expectations about the learning result believe that their discussion is comparatively not so efficient and reduces their motivation for collective learning (Zhai, 2021). Therefore, merely dividing students into several groups and telling them to learn in groups does not necessarily achieve high-quality learning or cooperation (Baloche and Brody, 2017). Teachers should consider students' concerns and difficulties. In this study, teachers need to help students improve their language and thinking abilities to increase their confidence to revise their writing. At the same time, they also need to consider making students engage in group work more actively.

11.2.3 Factors hindering thinking development

As discussed in section 2.2.1, the Chinese government has realised the significance of CT and explicitly put forward relevant policies and requests to conduct CT education in HEIs. The educational policies and CT requirements set a national context for students' CT development. Although it has remained at the policy level, and the integration of CT in different disciplines has not become commonplace (Zhang *et al*., 2020), it points out a general direction to stimulate teachers to conduct CT teaching practice. Guided by such policies and realising the significance, many Chinese EFL teachers strongly agree that CT should be taught in their classroom and have applied it to teaching practice, benefitting students' thinking. Even though the government and teachers have built a beneficial environment, other factors possibly hamper CT learning. For example, a teacher-centred pedagogy that stems from authority-respect Confucius culture. Students are supposed to be required to respect and be obedient to teachers in this learning context, especially when holding different opinions from teachers (Wan, 2001; Wang and Seepho, 2017; Tian and Low, 2011; Zhang *et al*., 2020; Zhong *et al*., 2021). Such studies indicate that individual students' behaviour or

performance in specific contexts can be profoundly influenced by a broad and general cultural backdrop or context (Tian and Low, 2011).

In addition, students' response to the learning context influences their behaviour and performance (Tian and Low, 2011). They usually adjust their learning methods or strategies to satisfy the requirements of a particular teacher, department or university (Clark and Gieve, 2006). Therefore, the traditional culture could not be the only barrier to Chinese students' thinking enhancement. As discussed in Chapter 1, section 1.3.4, although suffering a great shock when going abroad for the first time or not getting rid of the influence of Chinese traditional culture immediately, the Chinese international students make efforts to actively adapt to the CT-required western learning context. Their different attitudes or behaviour in the Chinese and western settings reveal that learning context profoundly influences students' thinking learning (Cheng, 2000; Clark and Gieve, 2006; Jones, 2005) in a pragmatic way (Tian and Low, 2011).

11.2.3.1 An assessment-oriented environment

Specific to this study, although a CT-enhanced learning context was created and students held a positive attitude, some factors within the learning context beyond teachers and students undermine CT teaching. Based on the result of students' suggestions for PBL class (section 9.3.3), the top three requests were about keeping and improving CT strategies for the next semester (n=21out of 88, 23.9%), adding CET4 teaching contents (n=20 out of 88, 22.7%) and increasing language knowledge teaching (n=18 out of 88, 20.5%). Students' suggestions on keeping CT strategies and acquiring more language knowledge indicated they recognised the significance of thinking and language teaching in this EFL course. The second one revealed their pragmatic or test-related English learning motivation for preparing and getting high scores in the CET4. As a prerequisite, the test result is directly related to undergraduates' "college graduation, graduate program admission, employment, promotion, and even residential permits" (Fan and Frost, 2022, p. 81). Given great opportunities or the high risks generated by the test result, test-takers are under high pressure to pass the test (Fan and Frost, 2022). As a result, they may fall into a potential pitfall of learning English for the test instead of meaningful language application capacities (Bai, 2020).

On the other hand, the university also faces the challenge of helping students achieve a higher passing rate in that examination as an index assessing the teaching quality and an essential instrument for the employment market. It means getting higher scores in the CET4 is no longer those test-takers' own business but the university and students. Although complaining about the extra burden of pursuing the passing rate (Fan and Frost, 2022), administrators have to build up such an environment to help students pass the examination. In this study, this assessment-oriented context is built by the School of Stomatology, where participants studied in this research. The new students were required to take the English mocks organised by the school almost every two weeks. The school faculties believed it would help students get accustomed to the question types in the CET4 as soon as possible. They even asked students to buy a vocabulary book and required them to memorise the words. Therefore, participants willingly or unwillingly had to prepare for the test at least one year in advance from the beginning of the first semester. What the school asked students to do is almost the same as the test-related learning experience in high school (section 10.2), misleading them to believe that their primary task in English learning in university is preparing for and passing that test. To some degree, the assessment-oriented context made by the university intensifies their test-taking pressure. The high pressure from themselves and the university makes them pay undue attention to the CET4. Furthermore, they may expect the English class in university to serve the same learning purpose as the high school to get high test scores. Therefore, it is no wonder they hoped teachers could provide test training.

However, such expectations contradicted the CE teaching objectives, such as cultivating language use abilities or critical thinking to prepare for future employment or higher-level academic studies (Chapter 1, section 1.3). In addition, focusing on teaching test-taking violates the prescription of the revised CECR: as an instrument for assessing students' English capacities, CET4/6 should not dominate or replace English teaching (The National Foreign Language Teaching Advisory, 2020, p.8). Meanwhile, students who merely engage in a massive amount of exercise may lose interest in English, especially those with higher language proficiencies. Dörnyei *et al*. (2006) and Bai (2020) argue that students are less motivated to learn English without learning achievement or fulfilment if they are forced to learn English with a sense of

obligation and duty or out of a utilitarian expectation from the test. In contrast, studies in western culture seldom discuss the influence of high-stakes examinations (Lee, 2018). In this sense, the assessment-oriented learning environment could hinder students' CT learning.

11.2.3.2 An absence of assessing CT in College English tests

The writing rubric of the CET4 is too vague for thinking requirements by simply stating whether test-takers present their topics clearly. At the same time, CT is not explicitly shown in the list of skills assessments in the CET exam syllabus (National College English Test committee, 2016). There is only one sentence implicitly listed in the language application scale: students should "apply cohesive devices to present the relationships between sentences" (p.3). The absent descriptions indicate that critical thinking is seriously underrated in this China's largest scale English test. This is inconsistent with the new teaching objective of the CE course (Chapter 1, section 1.2). As Tian (2008) and Liang and Fung (2021) claimed, traditional exams do not test thinking skills. As a result, students, especially those with test-related English learning motivations, may be passive about acquiring CT knowledge before taking part in the test because they believe it could not improve their test performance. An interviewee (S11) stated that he/she knew the CT importance, but the priority was to pass the CET4. So, CT learning should be postponed until passing that examination. That attitude may represent part of the participants' perspectives in this study. Their top priority is English language learning and test-taking skills. CT learning may hinder their English learning because their time is limited. This misunderstanding of ignoring the compatibility between English and thinking learning may lead to students' passive attitude when participating in collaborative work. I do not mean that these students should be criticised for their pragmatic consideration. My analysis only suggests that this assessment-oriented context and the absence of measuring thinking in examinations influence students' perspective or motivation to acquire thinking knowledge. For example, some students' inactive attitudes in group discussion (Chapter 10, section 10.4.2) or they paid more attention to correct linguistic mistakes than the thinking problems (Chapter 8, section 8.5.3). Therefore, practitioners may felt challenging to conduct thinking teaching in the assessment-oriented context. Moreover, the time allocation is limited for the CE course (section 1.3). Considering

the trade-offs, understanding students' current thinking state and what and how reasonably conduct CT teaching needs to be considered (Huber and Kuncel, 2016). Finally, another consideration is gaining authentic support from the university. After all, thinking development is an accumulation process over a long training period (Cáceres *et al*., 2020) that needs collaboration between teachers and students and contextual support (Zhang *et al*., 2020).

Unlike Tian (2008) claims that Chinese university teachers avoid teaching CT, there is an increasing number of teachers who have been conducting CT teaching in different disciplines at a current stage, particularly after the Chinese government's new round of curriculum reform (Wang and Seepho, 2017). However, CT skills are still not emphasised or explicitly required in exams (Tian, 2008), including the CET4. This deficiency may cause students to reduce their motivation to acquire CT knowledge under pressure to pass the examination.

11.2.3.3 Students' lack of application skills

There is a gap between my prediction of students' effective utilisation of the thinking checklist and their low percentage of using it to revise the essays. A thinking checklist was introduced in class as a valuable instrument for students to improve thinking skills in writing. Only six interviewees (out of 17, 35%) stated that their groups used it to revise their writing (Chapter 8, section 8.5.3), indicating six groups used that list since all the 17 interviewees came from 17 different groups. Nearly half of the interviewees (n=6 out of 14, 43%) who mainly corrected language mistakes stated that they felt it was difficult to apply the list to the draft revision. While only two (out of 14, 14%) mentioned that they misunderstood it. Other reasons were they forgot to use (n=3) or they thought it was time-consuming (n=3). In this light, lacking CT application skills seemed to be the greatest challenge from students' perspectives. In contrast, Zhang *et al*. (2020) reported that Chinese EFL learners' attitude towards CT is the major hindering factor from the teachers' perspective (n=263). The inconsistent findings demonstrated a potential gap between practitioners' and learners' perceptions when discussing challenges in CT learning. Although the sample sizes in current study and Zhang's study are quite different, the finding in this research may remind teacher-researchers that what teachers and students perceived could be divergent

owing to their different identities. The results would be more comprehensive if it is discussed from both teachers' and students' angles.

In addition, students' lack of application skills was also showed in their writing projects. The evidence from students' revision drafts showed that they still focus on language correction (Chapter 8, section 8.5.3) even though they stated that a good essay should demonstrate both correct language and critical thinking in the interview (Chapter 8, section 8.5.2). Considering the growing trend of CT awareness, students may need more time to be accustomed to the new standard and practice using the thinking checklist as an instrument to revise thinking in their essays. This inconsistency may further explain why students significantly increased the frequency of thinking skills in writing. However, the magnitude was modest.

Generally speaking, the CT teaching support from the government policy and the CT-enhanced teaching environment created by practitioners benefit learners' thinking development. However, the exam-oriented context created by the university may hamper CT learning inasmuch as the pressure to pass CET4, which is a specific adverse factor in the Chinese context. The beneficial and detrimental impacts caused by different contexts in this study confirm the significance of learning context for students' CT better performance (Jones, 2005). Cheng and Wan (2017) suggested that once determining to highlight a CT teaching objective in their courses, teachers should reflect on whether their classroom environment is consistent with the CT-enhanced learning context, such as an explicit teaching objective or CT strategies suggested in this study. Apart from external factors, the internal factor of novices' insufficient application skills also hindered their thinking development.

11.3 Students' CT Learning Outcomes

Having discussed a supportive context created in PBL teaching intervention and potential hindering factors, this section discusses students' CT performance outcomes. In particular, the inconsistent strength across scales of high Inquisitiveness, and low Truth-seeking and Open-mindedness and their consistent development across scales of generic and content-specific thinking skills. Students' positive attitudes and perspectives towards the effectiveness of CT strategies and PBL intervention have

showed the CT-oriented PBL approach has had an positive impact on their CT development.

11.3.1 CT performance

Generally, current research results reveal that CT-oriented PBL teaching intervention has helped to improve students' CT dispositions, generic thinking skills and content-specific thinking skills in writing. In terms of thinking disposition, students' increased awareness of using thinking skills in writing indicated their enhancement of a general mindset attribute towards CT. Specific to the dispositional dimensions, the mean scores of four out of six in the CCTDI increased. In terms of generic thinking skills, their mean scores of an Overall and all the three targeted core thinking skills developed (i.e. Inference, Analysis and Evaluation) in the CCTST. In terms of content-specific thinking skills, they significantly increased frequencies of the same targeted core thinking skills in writing.

11.3.1.1 Students' inconsistent strength across scales of high Inquisitiveness, and low Truth-seeking and Open-mindedness

Students significantly developed four out of six dispositional dimensions of CT disposition: Open-mindedness, Inquisitiveness, Analyticity and Confidence in Reasoning after PBL (Chapter 8, section 8.3.2). However, the increased scores kept the same score range, indicating their attitudes did not become more positive to a higher level. In particular, they held the same ambivalent attitude towards Open-mindedness, the same positive attitude towards Analyticity and Confidence in Reasoning and the same strong positive attitude towards Inquisitiveness before and after PBL (Chapter 8, section 8.3.1). Although not developing the Maturity in Judgement scale significantly (Chapter 8, section 8.3.2), participants kept a positive attitude before and after PBL (Chapter 8, section 8.3.1). While not increasing the mean score of the Truth-seeking scale obviously (Chapter 8, section 8.3.2), students kept an ambivalent attitude with the lowest mean score among all the six dispositions. This different growth magnitude pre-and-post-PBL indicated that CT-oriented PBL teaching intervention is practical for developing part of their mindset attributes. However, it failed to alter the unbalanced status of students' thinking disposition. Any

weakness for individual dimensions will adversely influence students' overall disposition and even lead to defective decisions (Insight Assessment, 2021).

The most positive attitude towards Inquisitiveness and ambivalent/inconsistent attitude towards Truth-seeking and Open-mindedness highlight students' thinking features in this study. A strong positive attitude of Inquisitiveness indicates that students have intense intellectual curiosity and are eager to acquire new knowledge that may not immediately or obviously be useful at the moment (Insight Assessment, 2021). Students' attitudes towards CT learning may explain their high Inquisitiveness. They learnt to draw thinking maps and participate in group discussions, even though they thought CT learning might not help them pass CET4 or was not directly related to their English learning at the current learning stage (Chapter 11, section 11.2.3.1). As for their suggestions for the PBL course in the questionnaire, some expressed their desire to explore further thinking knowledge. For instance, they hope for more logical training or fine-grained skills for drawing thinking maps (QS12, QS46).

Students' ambivalent/inconsistent attitudes towards Open-mindedness and Truth-seeking illustrated that they were not likely to show tolerance when acquiring knowledge. Or it seems they understood this newly acquired knowledge from their perspectives or stereotypes even though they had a strong desire to learn new things. In addition, they may neglect some relevant evidence or refuse to ask hard questions during their exploratory process.

In this study, although students realised the CT significance and changed evaluating standards in writing, they kept their previous standard of correcting linguistic errors as they did in high school rather than revising language and thinking. The gap between the new standard to evaluate writing and unchanged reviewing practice revealed that they seemed to accept and be aware of the necessity of CT when being taught such skills. However, some are less likely to use it when revising writing for multiple reasons, such as their thinking defects or limited revision time (Chapter 11, section 11.2.2.2). Their inconsistent behaviour reflected that students' attitudes, especially those who chose to follow the previous standard, were possibly passively open toward the new standard rather than investigate whether it benefited their better writing performance. In addition, students' weakness in truth-seeking is manifested in their giving up to use the thinking checklist when they did not know how to use it. Strong

critical thinkers should aim to approach problems instead of giving up intentionally or unintentionally (Insight Assessment, 2021), no matter whether they have adequate or inadequate abilities to solve problems or whether the problems are complex or not.

Therefore, students with strong Inquisitiveness but weak Open-mindedness and Truth-seeking generally meant they had a strong desire to acquire new knowledge. However, they still hold their preconceptions to interpret what they learnt. When meeting challenges, they may react with a negative attitude without courageous inquiry and integrity in order not to face these difficulties. Their general weaknesses observed in CT dispositions remind teachers to explore targeted training exercises to help students overcome such obstacles.

The result of students' inconsistent strength across dimensions of strong Inquisitiveness and weak Truth-seeking echoes the same findings of studies in the western context (Bixler *et al*., 2015; Colucciello, 1997; Noone and Seery, 2018) and the Chinese backdrops (Huang *et al*., 2021; Ip *et al*., 2000; Yeh and Chen, 2003; Tiwari *et al*., 2003; Zhang *et al*., 2009). It is reported that the mean scores in the above two dimensions for western undergraduates (i.e. in America and Australia) are higher than Chinese students (Dennett, 2014; Dennett and DeDonno, 2021; Ku and Ho, 2010; McBride *et al*., 2002; Tiwari, 2003). In comparison to the western culture, some studies (Gu and Schweisfurth, 2006; Ku and Ho, 2010; Paton, 2005; Shi, 2009; Tian and Low, 2011, Wang and Seepho, 2017) have attributed Chinese students' weaknesses in Truth-seeking to the traditional culture of pursuing consensus and avoiding arguments. It is assumed that teachers' higher authority in Chinese classrooms may make students, who are supposed to be at a lower status, not feel comfortable if question their received knowledge (Lucas, 2019).

Yet the tendency of high Inquisitiveness and low Truth-seeking seemed to be a universal characteristic for university students in both cultures based on the results of the aforementioned studies. With this in mind, this typical characteristic of Eastern and Western students cannot simply be attributed to culture. Lucas (2019) argued that some teachers neglect or downplay thinking elements or only teach a specific CT segment because of its complexity. However, few studies investigate why university students share this feature and corresponding teaching methods to overcome this obstacle.

The other feature of weak Open-mindedness in this study is seldom found in western students (Bixler *et al*., 2015; Colucciello, 1997; Ip *et al*., 2000; Noone and Seery, 2018). In other words, students in western culture rarely held an ambivalent, even a negative attitude towards that mindset attribute. Dennett and DeDonno (2021) point out that cultural context is a hindering factor that influences Chinese students' performance of Open-mindedness. Ku and Ho (2010) further elaborate that diverse opinions are valued and treated fairly in the west, while Chinese society may not appreciate diversity in opinion with respect for authority and social harmony. Besides the cultural factor, they report that Open-mindedness is positively related to Truth-seeking. Students are more likely to open their minds and actively engage in thinking in the context of being encouraged to pursue the truth. Given the close connection between the two dimensions and students' poor performance on both scales, teachers should create a class culture enabling students to really open their minds to accept peers' diverse suggestions, even negative comments and, at the same time, reminds them of making judgments based on the truth.

Based on the above discussion, culture seems to be a barrier to improving Truth-seeking and Open-mindedness in the Chinese context. Being aware of such deficiencies, educators could create a democratic learning environment and design tailor-made student-centred CT instructions to overcome weaknesses by encouraging students to ask questions (Dennett, 2014) or providing more opportunities to exchange their responses through group discussion and teamwork (Du *et al*., 2013).

11.3.1.2 Students' consistent strength across generic CT Skills and content-specific thinking skills

Generally speaking, students not only significantly developed the generic thinking skills in a standardised CT test (i.e. the CCTST) (Chapter 8, section 8.2) but also content-specific thinking skills in writing following a non-standardised measurement developed by myself (Chapter 8, section 8.4). Both developments indicated that CT-oriented PBL course design benefits students' thinking development. In addition, the magnitude of the Overall skill development in the CCTST (d=.57) and the total frequency of thinking skills in writing (d=.43) after PBL are moderate (Chapter 8, section 8.2.2). The medium magnitudes align with the moderate average effect size through a meta-analysis of 117 empirical studies by Abrami *et al*. (2008). While it is

not consistent with the research by Niu *et al*. (2013), claiming the degree of university students' thinking development is small after analysing 31 empirical research.

Furthermore, this study's generic CT skill outcome produces a higher effect size than the content-specific CT skills. This result is contrary to the study by Abrami *et al*. (2015), reporting that the effect size of content-specific skills is greater than the generic CT skill. The inconsistent results between the current study and Abrami *et al* imply that teaching interventions indeed make students develop significantly. However, the magnitudes of enhancement in generic and specific thinking skills are heterogeneous depending on different strategies.

Although the medium or modest magnitude of CT enhancement seems to be discouraging, it supports a perspective that it is difficult to foster thinking skills over a short period because thinking development is a gradual and accumulative process (Halpern, 2001; Niu *et al*., 2013; Cáceres *et al*., 2020). A longer teaching intervention (at least 12 weeks) is reported to be the crux of significant thinking development (Cáceres *et al*., 2020; Niu *et al*., 2013; Sun, 2017; Tsui, 2002). However, it does not mean the more prolonged the teaching period is, the more significant enhancement (Abrami *et al*., 2015). To solve such problems, practitioners could take full advantage of the prolonged teaching period by repeating CT strategies and spaced learning (Bixler *et al*., 2015) or extend the intervention scale by applying more courses (Cáceres *et al*., 2020). Therefore, suppose extending the teaching period may result in more remarkable development.

Thinking skills and dispositions are closely correlated and indispensable for ideal critical thinkers, who are supposed to be willing and able to think well (Chapter 2, section 2.1.1). It means if individuals display thinking skills even with higher levels but refuse to solve problems by using such skills unless heavily prompted, they are supposed not to be strong critical thinkers. Participants in this study have strong generic thinking skills but uneven performance among different dimensions of CT disposition, revealing that students could demonstrate sufficient thinking skills when solving problems. However, their weaknesses in Open-mindedness and Truth-seeking sometimes detrimentally impact their efforts toward a problem solution.

11.3.2 Attitudes and perceptions

When introducing a new CT pedagogy in the EFL course, it is essential to understand students' attitudes and perceptions based on the data collected from questionnaires and interviews. On the one hand, the effectiveness of CT strategies and PBL teaching affected participants' attitudes and perceptions. While on the other hand, their positive or negative attitudes and perceptions also impact their behaviour and performance in using such strategies. Respondents in this study who held positive attitudes towards CT strategies and PBL teaching were more likely to use and apply such skills in practice and may benefit more from this teaching intervention.

11.3.2.1 Students' attitudes towards and perceptions of CT strategies

In terms of CT skills, students held positive attitudes towards the effectiveness of three CT strategies (i.e., thinking maps, group discussion, and peer review) (Chapter 9, section 9.2.2.3) from the CT strategies questionnaire. Combining these strategies functioned as a complementation to support each other and improve different core thinking skills. In the face-to-face interview, they thought thinking maps developed their Interpretation, Analysis, and Inference skill (Chapter 10, section 10.3.1); group discussion enhanced their Evaluation and Explanation (Chapter 10, section 10.4.1), and peer review increased Analysis, Evaluation, and Self-regulation (Chapter 10, section 10.5.1). This enhancement covers all the core thinking skills in this study, and some are the skills the CCTST and writing did not demonstrate (Chapter 6, section 6.3.1, section 6.5.3.1), such as Self-regulation or Interpretation. Their statements further support the previous discussion of the feasibilities of integrating CT strategies into this PBL course design (Chapter 11, section 11.2.2.1). In addition, what they believed about the strategy's effectiveness may link to better performance in CT tests and increased frequencies of using thinking skills in writing after PBL. However, students' negative attitude also influenced their performance. Specific to this study, more than half of the groups did not use the thinking checklist to revise their writing during peer review, partly due to their inactive attitudes (Chapter 11, section 11.2.1.2). Their passivity may directly result in a small magnitude of grown frequencies of thinking skills in writing. In addition, some students' inactive attitudes towards participation in peer review and group discussion reduced the learning enthusiasm of active participants because of not receiving a response (Chapter 11, section 11.2.2.2).

The different learning outcomes resulting from learners' attitudes indicated that attitude could be an essential indicator.

In terms of CT disposition, students held the same positive attitudes towards the effectiveness of strategies to develop all dimensions of CT dispositions (Chapter 9, section 9.2.1.2). Their general active belief allows them to foster their CT awareness to use thinking skills and make them perform better in writing by increasing the frequencies significantly. Although they believed CT strategies contributed to the enhancement of thinking attributes and they indeed increased the mean scores of most dimensions in the CT test, students still held an ambivalent/inconsistent attitude towards weak performance in Truth-seeking and Open-mindedness (Chapter 11, section 11.3.1.1). This unsatisfied outcome revealed that CT strategies in this study might not be the targeted methods to change those two weak mindsets, or their effectiveness needs to be further proved, considering students' significantly development in many dispositions. Zhang *et al*. (2020) argue that Chinese traditional culture undermines students' disposition towards thinking besides the CT strategies (Chapter 11, section 11.3.1.1).

11.3.2.2 Students' attitudes towards and perceptions of PBL teaching

Students held positive attitudes towards the effectiveness of PBL in improving their CT, including teachers' scaffolding role, learners' awareness of the real world, and group work (Chapter 9, section 9.3.1). Their result means students believed these three core elements worked effectively in their thinking improvement. Teachers' acting role is verified by many studies (Alharbi, 2018; Hughes and Chen, 2011; Okolie *et al*., 2022), suggesting teachers in a student-centred classroom should act as active facilitators to help students acquire CT knowledge and encourage them to engage in activities collaboratively, in order to achieve the CT teaching objective finally. The emerging relationship between teachers and students has become a potent element in CT teaching and learning (Liberante, 2012).

In addition, students' positive attitudes towards the effectiveness of PBL in raising their awareness of the real world indicated that this teaching method motivates students to pay more attention to authentic problems. In this study, the diverse topics they learnt about university students' studies or lives may raise their motivation to

conduct further research in different directions within the same topic out of their interest. Students were given a chance to study their concerns in their studies or life by completing writing projects, contributing to thinking training, and their problem solutions (section 11.2.2.2).

Finally, students' positive attitude towards the effectiveness of group work revealed that they believed collective work was beneficial to thinking improvement. On the one hand, they could better prepare for their writing by brainstorming all the possible ideas (Chapter 10, section 10.4.1). While on the other hand, they receive much more constructive feedback through peer review (section 11.2.2.1).

The aforementioned three elements in PBL teaching are consistent with three categories of practical CT instructional approaches based on the study of Abrami *et al*. (2015). The first is *dialogue*, which emphasises discussion. The second, *authentic or anchored instruction*, refers to selecting real problems or situations out of students' interest. This is similar to the second element of learners' awareness of the real world in PBL teaching. At last, *mentoring*, a strategy of instruction between experts and novices, is equal to the first element of teachers' scaffolding role in PBL. Abrami *et al*. (2015) commented that combining the three categories achieves the best CT learning outcome. In this sense, PBL, featured with three elements in this study, could maximise its functions to gain a better CT result. Even though facing many challenges when studying collaboratively (section 11.2.2.2), students may become more accustomed to this new pedagogy as time goes by. Due to their inexperience, it is also the teachers' responsibility or challenge to help them familiarise themselves with CT learning (Zhang *et al*., 2020) and PBL teaching.

11.4 Summary

This chapter has answered the research questions by discussing the context created by the CT-oriented PBL course design and the hindering factors for thinking development and CT learning performance (i.e. thinking skills and disposition). It also discusses how CT strategies as a catalyst impact students' thinking development. Finally, it discusses the effectiveness of this course design by investigating learners' attitudes and perspectives. The CT-oriented learning context and students' positive CT learning outcomes and attitudes imply the PBL's effectiveness to improve thinking in

the Chinese context and the feasibility of integrating CT teaching into an EFL course. The positive result in this study is consistent with several studies (Cáceres *et al*., 2020; Davis, 2006; McPeck, 1981; Moore, 2015), claiming CT could be taught within a specific subject. Although Abrami *et al*. (2008) hold a different perspective to claim that it is less critical whether CT is taught separately from the content or embedded with it, they recognise the significance of explicit CT requirements as an essential part of the course design. This perspective is also verified in this study and more research (Snider, 2017; Soufi and See, 2019; Sun, 2017; Tsui, 2002) that explicit teaching is an effective way to achieve better CT outcomes.

In this part, the first three chapters provided three main findings in accordance with three research questions. Chapter 8 reported CT learning outcomes. Student enhanced their overall and all three targeted core thinking skills in CT tests and writing. In comparison, their CT disposition developed unevenly. Chapter 9 revealed participants' positive attitudes towards CT strategies and PBL elements. Moreover, thinking maps was rated the most effective method. Chapter 10 explained how CT strategies impact their thinking development by displaying the development of six core thinking skills, especially the Analysis and Evaluation, and their concerns when using strategies, such as their restricted competence in English and thinking and the low efficiency of the group work. Chapter 11 discussed the findings in two strands. Firstly, it discussed the context for CT enhancement based on the CT-oriented PBL course design and hindering factors. Secondly, it discussed students' CT learning outcomes by elaborating on their CT performance and their attitudes and perceptions.

PART IV CONCLUSION

Chapter 12 Conclusion

12.1 Introduction

On the basis of the previous part of the research findings and discussions, this conclusion chapter briefly reviews the aims and main findings to answer the research questions and present the contributions, implications, and limitations. Finally, the recommendations for future research are illustrated.

12.2 Summary of Main Findings

12.2.1 CT development university students made after PBL intervention

The impact of PBL teaching intervention on students' thinking manifested in two dimensions, including thinking dispositions and skills. Regarding CT dispositions, the CCTDI pre-and-post PBL revealed that students increased most dimensions: Inquisitiveness, Analyticity, Open-mindedness and Confidence in Reasoning. However, students did not change their unbalanced strength across scales of high Inquisitiveness and low Truth-seeking and Open-mindedness. The interview data revealed students generally increased their CT awareness by transforming writing standards from focusing on language and handwriting to language and critical thinking.

Regarding thinking skills, the CCTST before and after PBL presented that students significantly improved their generic thinking skills on the Overall scale and all the three targeted core thinking skills (i.e. Inference, Analysis and Evaluation). This result is confirmed by students' content-specific thinking performance in writing: they increased frequencies of using thinking skills overall and of the same core thinking skills in writing. In addition, the CCTST also suggested their Overall thinking skill was at a strong level before and after PBL, which means they are skilled in using reasoning to make judgements or solve problems.

As discussed in Chapter 2, section 2.1.1 and Chapter 11, section 11.3.1, strong critical thinkers should employ consistent internal motivations to make decisions or solve problems using critical thinking skills. To their performance in the above CT tests, students in this study have adequate thinking skills to make decisions or solve

problems. At the same time, they are improving at using more thinking skills in their writing. However, they do not always accept different opinions, perspectives or suggestions and are possibly ready to give up when meeting some difficulties.

12.2.2 Students' attitudes towards and perspectives of CT strategies and PBL teaching intervention

The questionnaire suggested students hold positive attitudes towards the effectiveness of CT strategies and PBL classes. They kept the most positive attitudes towards thinking maps. The interview data confirmed their attitudes, further demonstrating students thought it was the most helpful strategy. In comparison, peer review was students' favourite strategy. Students held the same positive attitudes towards three PBL elements: a teacher's scaffolding role, students' awareness of the authentic world, and their group work. There is no significant difference among them, indicating that students believed these elements contribute equally to their thinking learning.

Although they are generally optimistic about the CT strategy and PBL teaching, interview data revealed that some students present their passivity when using the thinking checklist to revise essays (Chapter 11, section 11.2.1.2) and taking part in collaborative work (Chapter 11, section 11.2.2.2). The passive attitudes undermined their performance in such activities, such as limited achievement in peer review and their low efficiency of not taking full advantage of discussion time (Chapter 10, section 10.4.2). Besides a common reason for Chinese culture's negative influence, two causes stem from an assessment-oriented environment (Chapter 11, section 11.2.3.1) and an absence of assessing CT in tests (section 11.2.3.2).

12.2.3 The impact of CT strategies on students' CT improvement

The interview data demonstrated that the selected language teaching methods could be CT strategies to improve students' different core thinking skills. The strategy of thinking maps improved their Interpretation, Analysis and Inference (Chapter 10, section 10.3.1); group discussion benefited their Evaluation and Explanation (Chapter 10, section 10.4.1), and peer review contributed to their Analysis, Evaluation and Self-regulation (Chapter 10, section 10.5.1). As a supplement, these results elaborated on what students reflected about the effectiveness of such strategies to improve

thinking skills besides the standardised test of CCTST (Chapter 8, section 8.2) and the writing analysis test (Chapter 8, section 8.4).

The data from the interview also presented students' concerns, difficulties or dissatisfaction when engaging in these strategies. Their weaknesses mainly lie in their insufficient abilities to detect thinking defects when discussing thinking maps and peer review contents. Their primary dissatisfaction was the low efficiency and unequal participation in group work.

12.3 Contributions

The current study's value lies in integrating critical thinking education into PBL teaching classes, aiming to enhance students' critical thinking in tertiary education. The present EFL teaching in China is far from satisfactory in strengthening students' CT capacities, including thinking teaching kept at the policy level without clear guidance (Chapter 1, section 1.4) and students' poor thinking performance in a teacher-dominated class (section 1.3.4). Hence, this research provides teachers in university with a student-centred and CT-oriented PBL course design (Figure 4.1). For the first time, this research project employed three language teaching methods as CT strategies deepened student learning in their academic study and provided comprehensive perspectives to solve problems in their daily life through completing different writing projects.

In terms of PBL teaching, this study is one of the few PBL empirical studies with a precise CT teaching aim on the EFL subjects in tertiary education. The effectiveness of embedded CT strategies provided feasible options for practitioners to employ in PBL teaching. The current study may overcome a salient obstacle in EFL project-based learning: what activities can be effectively applied to help students communicate or cooperate in PBL (Chapter 3, section 3.3.3). Furthermore, current PBL teaching studies mainly concentrate on the overall description of CT development instead of exploring particular skills (Chapter 1, section 1.4). This overall thinking enhancement depicts a full-length portrait of students perceiving their thinking change in general rather than recognising which thinking skills they need to improve. In this light, they may lose a direction they could make efforts. Meanwhile, teachers could not utilise targeted teaching strategies to overcome learners' thinking

weaknesses. Therefore, this study contributes to providing a possibility of what exact thinking skills or dispositions PBL could develop, although different CT strategy selections may result in different thinking outcomes within the PBL teaching.

Furthermore, the current study provided a fine-grained pedagogical PBL framework concentrating on CT development specific to the Chinese background. Firstly, it defined concrete CT skills for first-grade undergraduates based on the writing requirements of the CSE 5 (NEEA, 2018) and CT descriptions of the Delphi Report (Facione, 1990). This particular CT requirement makes it feasible for Chinese EFL teachers to have a targeted teaching direction and for students to use thinking skills in writing and revising essays. This method of defining CT skills in writing could also contribute to defining CT requirements in other language disciplines, such as reading, listening and speaking. Once accomplishing all these subjects of language comprehension (i.e. listening and reading) and language expression (i.e. speaking and writing), EFL teachers may have a comprehensive CT understanding of the EFL discipline. If teachers apply their CT teaching to the above four courses, students will benefit from more thinking training exposure.

Secondly, this PBL research provides an option of student-centred pedagogy by allowing students to choose any title they are interested in and determine what they should discuss according to their different challenges or concerns. Students become more active in the PBL classroom than in the traditional classes. Exploration of PBL teaching that features students-centred provides a model for other practitioners accustomed to a conventional teaching method and intend to engage students in study actively. Meanwhile, this study echoes the national requirement of educational reform in China that encourages teachers to utilise multiple pedagogies, such as task-based, project-based, cooperation-based or inquiry-based teaching methods in the classrooms, to encourage students to transform from passive to active learning. Therefore, this study confirms the notion that a teaching intervention design that is compatible with the context could drive a better teaching outcome.

In terms of CT teaching, this research supports the perspective that CT training could be embedded in the EFL subject content knowledge teaching (Chapter 2, section 2.2.3.3) for students' better CT performance, especially in thinking skills. Additionally, positive CT learning outcomes by using language teaching methods as CT-enhanced

strategies also revealed that practitioners could explore some prevalent teaching methods or course elements in their subjects to conduct CT education. And this study is a workable demonstration of how these prevalent language teaching strategies can be employed to enhance language learners' thinking skills.

Specific to the Chinese context, as discussed in Chapter 1, section 1.3.1, although educational policymakers and researchers have been appealing to integrate CT into disciplinary fields, it has not become commonplace. Furthermore, most current CT studies in EFL courses are for English majors. This study addresses the gap in research by providing an empirical study of embedding CT instruction in a systematical pedagogy to teach CT for non-English majors. Finally, this study provides a clear and specific definition of CT skills in Chinese EFL writing, which helps other teachers teach and assess thinking skills.

12.4 Implications for Practice

Integrating CT teaching into an EFL class employs an inherent advantage owing to a close relationship between language and critical thinking. As a student-centred pedagogy, PBL teaching intervention is expected to allow students to apply CT knowledge acquired in class to complete projects and eventually contribute to CT development. However, merely PBL teaching without explicit CT instructions may run a risk of achieving satisfactory outcomes. This study attempts to explicitly integrate CT teaching into PBL pedagogy in a Chinese EFL course (i.e. the CE course) and investigate whether it improves students' thinking skills and raises their CT awareness. Specifically, this study revises a ten-step PBL framework in an L2 classroom designed by Alan and Stoller (2005) by infusing CT teaching within the framework to maximise PBL's effect on thinking and refining teachers' scaffolding role considering students' inexperience with thinking training and PBL learning. Students' better CT performance and positive attitudes reveal that this infusion approach is applicable and practical for thinking development in EFL high education.

In addition, most current PBL studies claim that it improves generic critical thinking but fails to elaborate further on the particular thinking skills. This deficiency may not allow students to detect their thinking weaknesses, resulting in a slim chance of tackling difficulties. Therefore, it would be better if there is any teaching strategy that

could help students develop core thinking skills and, at the same time, benefit their project completion. Methods of thinking maps, peer review and group discussion were custom-made for the form of projects in this study: essay writing. Meanwhile, as universal language teaching strategies, they are proved to be applicable and practical to thinking teaching, according to the results of this study. This success implies that different subject practitioners could use their common instructional approaches to thinking education, which is a practical option for teachers, especially those who did not receive CT professional training (Tsui, 2002).

Besides, a rubric with three thinking skills categories (i.e. Analysis, Inference and Evaluation) and ten corresponding items enable teachers to conduct thinking instructions in a targeted direction taking account of an absence of thinking skills teaching guidance in a Chinese EFL context. At the same time, students were given a more detailed thinking checklist to complete writing projects and revise peers' essays with higher quality.

It is universally believed that the traditional Chinese culture of respecting authority negatively influences students' thinking cultivation and development. However, it is not the only reason. In this study, an exam-oriented environment that requires students to pass the CET4 is an unexpected major hindering factor because it directly results in students' high pressure to prepare for the exam. In addition, the CET4 does not explicitly test thinking. As a result, some students may reduce their motivation to learn CT and still focus mainly on language learning. Because of the significance of linguistic proficiency for critical thinking and English tests, how to balance thinking and language teaching is another issue teachers should take into consideration in the context of Chinese high education.

12.5 Limitations

The main limitation of the current research is the absence of a control group. Due to the university's arrangement where I conducted this study, I taught two classes. To meet the statistical sample size, I chose both classes (n= 94), ensuring enough numbers of participants to take part in the study from the statistical consideration. Once the control and experimental group are involved in this research, it may risk a small sample size for valid statistical data analysis. In addition, this is an exploratory

study to assess the applicability of this CT-oriented PBL teaching. The research purpose could be achieved if the same students could develop their thinking after PBL intervention. However, the results of this study should be carefully considered concerning the effectiveness of CT strategies in PBL teaching since no control group was available to compare and prove these positive results were solely a result of using strategies in the instruction. This weakness is a typical limitation in so much CT teaching research in EFL (Soufi and See, 2019). Whereas results from different sources of CT tests, students' writing projects, questionnaires and interviews provide valid evidence to triangulate each other. Therefore, this study could provide evidence of the effectiveness of integrating CT in PBL teaching to improve students' thinking.

Next, the simplified Chinese version of the CCTST does not assess all six core thinkin skills. In order to triangulate the test result, the writing analysis also stressed the same core thinking skills. Therefore, this study did not evaluate students' CT performance in the other three skills of Interpretation, Explanation and Self-regulation. As mentioned in Chapter 6, section 6.3.1 and section 6.5.3.1, although the English version of the CCTST assessed five core thinking skills except for Self-regulation, I chose the simplified Chinese version to eliminate a hindering factor that may impact these Chinese participants' CT test performance considering their English proficiencies. In addition, even using the English version, the CCTST cannot test all the six skills mentioned before. This study then used interviews to obtain detailed information about how CT strategies improved thinking skills, which partly complements students' thinking skills' development.

The final limitation is that the duration of the PBL teaching intervention. Although the intervention lasted for an entire academic semester (15 weeks), it is only 20 hours in total, which was insufficient to teach students who had never trained in CT. Although 12 weeks have been regarded as the minimum time span for students' significant CT development, and the treatment length is not a crucial factor for a more excellent teaching outcome (Chapter 11, section 11.3.1.2), a more extended period with more effective thinking training is expected to help students lead into a more significant enhancement in the gradual and cumulative process. Considering the growing trend of CT awareness, students may need more time to be accustomed to the new standard

and practice using the thinking checklist as an instrument to revise thinking in their essays.

12.6 Recommendations

Based on Discussion Chapter 11 and the limitations mentioned above, several directions for future research are suggested as follows:

For researchers, they could design research with a control and experimental group merely using PBL teaching, allowing them to examine whether positive results can result from integrating CT instruction in the PBL teaching intervention alone. In addition, they could create an influential learning community for practitioners to acquire more CT professional knowledge, increase their awareness and improve CT instructional literacy. Specific to the Chinese context, they could investigate the factors that impact students' Truth-seeking and Open-mindedness and explore the targeted teaching methods to develop such mindset attributes. In this way, students could overcome their weaknesses in disposition and thinking skills and achieve more significant development. Once achieved, students will be more motivated to engage in thinking learning.

For practitioners, they could investigate and examine practical teaching approaches to minimise the detrimental influence of traditional culture and prompt students to participate more actively in CT activities individually and collectively. Specific to the Chinese culture, it is suggested to research how to build up a more democratic class environment to encourage students to ask questions and create a compelling interactive channel between teachers and students. EFL teachers could define different CT requirements in other subjects, like reading, speaking and listening classes, and further establish and investigate a comprehensive framework of connecting language and critical thinking teaching in a stereoscopic dimension. They also could examine the effectiveness of the CT-oriented PBL teaching intervention framework in this study in other educative practices with a similar cultural background, like other East Asian countries. Future studies could increase treatment duration, for instance, at least one academic year, so as to achieve a better teaching outcome. Since specific thinking skills may cover many items, it would not have

enough time to teach students all the detailed literature, or students may not have sufficient time to practice in a short period.

Policymakers or leaders in HEIs need to adapt to the exam-oriented environment for students to reduce the pressure of the test results and have more time or chances to acquire CT knowledge. They also need to increase the percentage of testing CT in the examinations and provide more opportunities for students to exercise their thinking skills in tests.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R. and Zhang, D. (2008) Instructional interventions affecting critical thinking skills and dispositions: a stage 1 meta-analysis. *Review of Educational Research*. **78(**4), pp.1102-1134.

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A. and Persson, T. (2015) Strategies for teaching students to think critically: a meta-analysis. *Review of Educational Research*. **85**(2), pp.275–314.

Afshar, H. S., Movassagh, H. and Arbabi, H. R. (2017) The interrelationship among critical thinking, writing an argumentative essay in an L2 and their sub-skills. *The Language Learning Journal*. **45**(4), pp.419-433.

Alan, B., and Stoller, F. L. (2005) Maximizing the benefits of project work in foreign language classrooms. *English Teaching Forum*. **43**(4), pp. 10-21.

Alharbi, M. A. (2018) Patterns of EFL learners' and instructors' interactions in asynchronous group discussions on free writing. *Journal of Information Technology Education*. **17**, pp.505–526.

AlKhoudary, Y. A. M. (2015) The effect of teaching critical thinking on Al-Buraimi University college students' writing skills: a case study. *International Journal of Applied Linguistics and English Literature*. **4**(6), pp.212-219.

Alnofaie, H. (2013) A framework for implementing critical thinking as a language pedagogy in EFL preparatory programs. *Thinking Skills and Creativity*. **10**, pp.154–158.

Al-Shaer, I. M. R. (2014) Employing concept mapping as a pre-writing strategy to help EFL learners better generate argumentative compositions. *International Journal for the Scholarship of Teaching and Learning*. **8**(2), pp.1-32.

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

Báez, C. P. (2004) Critical thinking in the EFL classroom: the search for a pedagogical alternative to improve English learning. *Ikala: Revista de Lenguaje y Cultura.* **9**(15), pp.45–80.

Bağ, H. K. and Gürsoy, E. (2021) The effect of critical thinking embedded English course design to the improvement of critical thinking skills of secondary school learners. *Thinking Skills and Creativity*. **41**, pp.1-13.

Bagheri, F. (2015) The relationship between critical thinking and language learning strategies of EFL learners. *Journal of Language Teaching and Researching.* **6**(5), pp.971-975.

Bai, Y. (2020) The relationship of test takers' learning motivation, attitudes towards the actual test use and test performance of the College English Test in China. *Language Testing in Asia*. **10**(1), pp.1–18.

Baloche, L. and Brody, C. M. (2017) Cooperative learning: exploring challenges, crafting innovations. *Journal of Education for Teaching*. **43**(3), pp. 274-283.

Bas, G. (2011) Investigating the effects of project-based learning on students' academic achievement and attitudes towards English lesson. *The Online Journal of New Horizons in Education* [online]. **1**(4), pp.1-25. Available from: https://scholar.google.com/scholar?hl=zh-CN&as_sdt=0%2C5&q=Investigating+the+Effects+of+Project-Based+Learning+on+Students%E2%80%99+Academic+Achievement+and+Attitudes+towards+English+Lesson.+&btnG= [Accessed 28th May 2019].

Beckett, G. H., (1999) Project-based instruction in a Canadian secondary school's ESL classes: goals and evaluations. Ph.D. University of British Columbia.

Beckett, G. H. (2002) Teacher and student evaluations of project-based instruction. *TESL Canada Journal*. **19**(2), pp.52-66.

Beckett, G. H. and Slater, T. (2005) The project framework: a tool for language, content, and skills integration. *ELT Journal*. **59**(2), pp.108-116.

Beckett, G. H. and Slater, T. (2018a) Project-based learning and technology. In: Liontas, J. I. (ed.) *The TESOL encyclopedia of English language teaching*. Oxford: Wiley Blackwell, pp.1-7.

Beckett, G. H. and Slater, T. (2018b) Technology-integrated project-based language learning. In: Chapelle, C. A. (ed.) *The concise encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell, pp.1-8.

Beiki, M., Gharagozloo, N. and Raissi, R. (2020) The effect of structured versus unstructured collaborative pre-writing task on writing skills of the Iranian EFL students. *Asian-Pacific Journal of Second and Foreign Language Education*. **5**(1), pp.1–29.

Bensley, D. A. (1998) *Critical thinking in psychology: a unified skills approach*. Belmont: Thomson Brooks.

Benson, P. (2005) *Autonomy in language teaching*. Beijing: Beijing Foreign Language Teaching and Research Press.

Bergman, M. M. (2011) The good, the bad, and the ugly in mixed methods research and design. *Journal of Mixed Methods Research*. **5**(4), pp.271-275.

Bernstein, B. (1974) Sociology and the sociology of education: a brief account. In: Rex, J. (ed.) *Approaches to sociology: an introduction to major trends in British sociology.* London: Routledge and Kegan Paul, pp.145-159.

Biesta, G. (2012) Mixed methods. In: Arthur, J., Waring, M., Coe, R. and Hedges, L. V. (eds.) *Research methods and methodologies in education.* London: Sage, pp.147-152.

Biggs, J. B. (1996) Western misperceptions of the confucian-heritage learning culture. In: Watkins, D. A. and Biggs, J. B. (eds.) *The Chinese learner: cultural, psychological and contextual influences*. Hong Kong and Melbourne: Comparative Educational Research Centre and Australian Council for Educational Research Ltd, pp.45–67.

Bilik, Özlem, Kankaya, E. A. and Deveci, Z. (2020) Effects of web-based concept mapping education on students' concept mapping and critical thinking skills: a double blind, randomized, controlled study. *Nurse Education Today*. **86**, pp.1-6.

Bilsborough, K. (2013) TBL and PBL: two learner-centred approaches. *British Council* [online]. Available from: https://www/teachingenglish.org.uk/article/tbl-pbl-two-learner-centred-approaches [Accessed 20 July 2022].

Bixler, G. M., Brown, A., Way, D., Ledford, C. and Mahan, J. D. (2015) Collaborative concept mapping and critical thinking in fourth-year medical students. *Clinical Pediatrics*. **54**(9), pp.833–839.

Bloom, B. S. (1956) *Taxonomy of educational objectives: the classification of educational goals. Handbook 1: cognitive domain*. New York: Longmans, Green and Co LTD.

Bloor, M. (2011) Techniques of validation in qualitative research: a critical commentary. In: Miller, G. and Dingwall, R. (eds.) *Context and method in qualitative research*. London: SAGE Publications Ltd, pp.38-50.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W. and Soloway, E. (2000) Creating usable innovations in systemic reform: scaling up technology-embedded project-based science in urban schools. *Educational Psychologist*, **35**, pp.149–164.

British Council (2020) Academic mean performance by nationality. *British Council* [online]. Available from: https://www.ielts.org/for-researchers/test-statistics/test-taker-performance [Accessed 19th Feb 2022].

British Council (2022) What is CSE? *British Council* [online]. Available from: https://www.britishcouncil.cn/en/exams/aptis/cse#:~:text=What%20is%20CSE%3F,%2C%20speaking%2C%20reading%20and%20writing [Accessed 4th Sep 2022].

British Educational Research Association [BERA] (2018) *Ethical guidelines for educational research* [online]. 4th edition. Available from: https://www.bera.ac.uk/wp-content/uploads/2018/06/BERA-Ethical-Guidelines-for-Educational-Research_4thEdn_2018.pdf [Accessed 24th May 2022].

Brinkmann, S. and Kvale, S. (2019) *Doing Interview*. London: SAGE Publications Ltd.

Brookfield, S. D. and Preskill, S. (1999) *Discussion as a way of teaching: tools and techniques for democratic classrooms*. San Francisco: Jossey-Bass.

Brophy, J. and Alleman, J. (1991) Activities as instructional tools: a framework for analysis and evaluation. *Educational Researcher*. **20**(4), pp.9-23.

Brumfit, C. (1984) *Communicative methodology in language teaching.* London: Cambridge University Press.

Bryman, A. (2016) *Social research methods*. 5th ed. Oxford: Oxford University Press.

Butchart, S. and Forster, D. (2009) Improving critical thinking using web based argument mapping exercises with automated feedback. *Australasian Journal of Educational Technology*. **25**(2), pp.268–291.

Cáceres, M., Nussbaum, M. and Ortiz, J. (2020) Integrating critical thinking into the classroom: a teacher's perspective. *Thinking Skills and Creativity*. **37**. pp.1-18.

Cali, K. (2021) Organisation. *Luwen Learning* [online]. Available from: https://courses.lumenlearning.com/basicreadingandwriting/chapter/outcome-organizing/ [Accessed 20 November 2021].

Carey, M. (2012) *Qualitative research skills for social work: theory and practice*. Abingdon: Routledge.

Chainklin, S. (2003) The zone of proximal development in Vygotsky's analysis of learning and instruction. In: Kozulin, A., Gindis, B., Ageyev, V. S. and Miller, S. M. (eds.). *Vygotsky's educational theory in cultural context*. New York: Cambridge University Press, pp.39-64.

Chang, J. Y., Mi, W. and Zeng, X. H. (2019) Impacts of project-based learning of the SEC courses on English majors. *Foreign Languages in China*. **16**(1), pp. 55-61.

Chen, J. and Zhang, L. J. (2019) Assessing student-writers' self-efficacy beliefs about text revision in EFL writing. *Assessing Writing*, **40**, pp. 27–41.

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

Chen, S.-L., Liang, T., Lee, M. L. and Liao, I-C. (2011) Effects of concept map teaching on students' critical thinking and approach to learning and studying. *The Journal of Nursing Education*. **50**(8), pp.466–469.

Chen, Y. and Zhang, C. (2019) Reform and opportunities: China English in Chinese higher education. In: Fang, F. and Widodo, H. P. (eds.). *Critical perspectives on global Englishes in Asia: language policy, curriculum, pedagogy and assessment.* Bristol: Blue Ridge Summit, pp. 176-193.

Cheng, M. H. M. and Wan, Z. H. (2017) Exploring the effects of classroom learning environment on critical thinking skills and disposition: a study of Hong Kong 12th graders in Liberal Studies. *Thinking Skills and Creativity.* **24**, pp.152–163.

Cheng, X. (2000) Asian students' reticence revisited. *System*. **28**(3), pp.435–446.

Cheung, C., Rudowicz, E., Kwan, A. S. F. and Yue, X. D. (2002) Assessing university students' general and specific critical thinking. *College Student Journal.* **36**(4), pp.504–526.

Cho, H. (2017) Synchronous web-based collaborative writing: factors mediating interaction among second-language writers. *Journal of Second Language Writing*. **36**, pp.37–51.

Clark, R. and Gieve, S. N. (2006) On the discursive construction of 'the Chinese learner'. *Language, Culture and Curriculum*. **19**(1), pp.54–73.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences.* 2nd ed. London: Lawrence Erlbaum Associates.

Cohen, L., Manion, L. and Morrison, K. (2018) *Research Methods in Education.* 8th ed. New York: Routledge.

Cole, M. and Gajdamaschko, N. (2007) Vygotsky and culture. In: Daniels, H., Cole, M. and Wertsch, J. V. (eds.). *The Cambridge companion to Vygotsky.* Cambridge: Cambridge University Press, pp. 193-211.

Colucciello, M. L. (1997) Critical thinking skills and dispositions of baccalaureate nursing students: a conceptual model of evaluation. *J. Prof. Nurs*. **13**(4), pp.236–245.

Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., Drohojowska, S., Saco, L. and Nelson, E. (2017) Project-based learning: a literature review. *MDRC* [online].

Available from: https://www.mdrc.org/sites/defalut/files/Project-Based_Learning-LitRev_Final.pdf [Accessed 22 May 2019].

Cooper, J. L. (1995) Cooperative Learning and Critical Thinking. *Teaching of Psychology*, **22**(1), pp. 7-9.

Cortazzi, M. and Jin, L. (1996) Cultures of learning: language classrooms in China. In: Coleman, H. (ed.), *Society and the Language Classroom*. Cambridge University Press, Cambridge, pp.169–206.

Cottrell, S. (2017) *Critical thinking skills: effective analysis, argument and reflection*. 3rd ed. London: PALGRAVE.

Creswell, J. W. (2007) *Qualitative inquiry and research design: choosing among five approaches*. 2nd ed. Thousand Oaks: Sage Publications.

Creswell, J. W. and Guetterman, T. C. (2018) *Educational research: planning, conducting, and evaluating quantitative and qualititative research*. 6th ed. New Jersey: Pearson Education.

Dale, D. (1991) An empirical taxonomy of critical thinking. *Journal of Instructional Psychology*. **18**(2), pp.79–92.

D'Alessio, F. A., Avolio, B. E. and Charles, V. (2019) Studying the impact of critical thinking on the academic performance of executive MBA students. *Thinking Skills and Creativity*. **31**, pp. 275–283.

Daley, B. J. and Torre, D. M. (2010) Concept maps in medical education: an analytical literature review. *Medical Education*. **44**(5), pp.440–448.

Dallimore, E. J., Hertenstein, J. H. and Platt, M. B. (2004) Classroom participation and discussion effectiveness: student-generated strategies. *Communication Education*. **53**(1), pp.103-115.

D'Antoni, A. V., Zipp, G. P., Olson, V. G. and Cahill, T. F. (2010) Does the mind map learning strategy facilitate information retrieval and critical thinking in medical students? *BMC Medical Education*. **10**(1), pp.61-71.

Darani, H. L. and Nafiseh, H. (2019) Pedagogical utility of oral discussion versus collaborative drafting. *English Teaching: Practice and Critique*. **18**(4), pp.464–477.

David, J. L. (2008) Project-based learning. *Education Leadership*. **65**(5), pp. 80-82.

Davidson, B. W. (1998) Comments on Dwight Atkinson's 'a critical approach to critical thinking in TESOL': a case for critical thinking in the English language classroom. *TESOL Quarterly*. **32**(1), pp.119–123.

Deng, Y. and Wang, X. (2009) An empirical study on developing learner autonomy through project-based learning. *Foreign Languages and Their Teaching*. **8**, pp. 31-34.

Dennett, S. K. (2014) A study to compare the critical thinking dispositions between Chinese and American college students. Ph. D. Florida Atlantic University.

Dennett, S. K. and DeDonno, M. A. (2021) A comparison between Chinese and American male and female college students' critical thinking dispositions. *International Journal of Chinese Education*. **10**(3), pp.1-14.

Denscombe. M. (2017) *The good research guide: for small-scale social research projects*. 6th ed. London: Open University Press.

Dewaelsche, S. A. (2015) Critical thinking, questioning and student engagement in Korean university English courses. *Linguistics and Education*. **32**, pp.131–147.

Dewey, J. (1933) *How we think: a restatement of the relation of reflective thinking to the educational process*. 2nd ed. Lexington, MA: D.C. Heath and company.

Deng, Z. (2014) Confucianism, modernization and Chinese pedagogy: An introduction. *Journal of Curriculum Studies*, **43**(5), pp.561-568.

Djiwandono, P. I. (2013) Critical thinking skills for language students. *TEFLIN Journal*, **24**(1), pp.32–47.

Dong, Y. N. (2017) Teaching and assessing critical thinking in second language writing: an infusion approach. *Chinese Journal of Applied Linguistics,* **40**(4), pp.431-450.

Dörnyei, Z. (2007) *Research methods on applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Dörnyei, Z. (2009) *Questionnaires in second language research: construction, administration, and processing*. 2nd ed. New York: Routledge.

Dörnyei, Z. and Csizér, K. (2012) How to design and analyze surveys in second language acquisition research. In: Mackey, A. and Gass, S. M. (eds.) *Research methods in second language acquisition: a practical guide*. New York: Wiley-Blackwell, pp.74-94.

Ding, X. W. (2016) The effect of WeChat-assisted problem-based learning on the critical thinking disposition of EFL learners. *International Journal of Emerging Technologies in Learning (iJET)*. **11**(12), pp.23–29.

Du, X. S. and Zhang, L. (2022) Investigating EFL learners' perceptions of critical thinking learning affordances: voices from Chinese university English majors. *SAGE Open*. **12**(2), pp.1-12.

Du, X. Y., Emmersen, J., Toft, E. and Sun, B. Z. (2013) PBL and critical thinking disposition in Chinese medical students-A randomized cross-sectional study. *Journal of Problem Based Learning in Higher Education*. **1**(1), pp.72-83.

Duff, P. A. (2006) Beyond generalizability: contextualization, complexity, and credibility in applied linguistics research. In: Chalhoub-Deville, M., Chapelle, C. A. and Duff, P. A. (eds.) *Inference and generalizability in applied linguistics: multiple perspectives*. Philadelphia: John Benjamins Publishing, pp.66-95.

Duff, P. A. (2012) How to carry out case study research. In: Mackey, A. and Gass, S. M. (eds.) *Research Methods in second language acquisition: a practical guide*. New York: Wiley-Blackwell, pp.95-116.

Duff, P. A. (2014) Case study research on language learning and use. *Annual Review of Applied Linguistics*. **34**, pp.233-255.

Duff, P. A. (2020) Case study research: making language learning complexities visible. In: McKinley J. and Rose H. (eds.) *The Routledge handbook of research methods in applied linguistics*. New York: Routledge, pp.144-153.

Duff, P. A. and Anderson, T. (2016) Case study research. In: Brown, J. D. and Coombs, C. (eds.) *Cambridge guide to language research*. Cambridge: Cambridge University Press, pp.112-118.

Dunlea, J., Spiby, R., Wu, S., Zhang, J. and Cheng, M. (2019) *China's standards of English language ability (CSE): linking UK exams to the CSE* [online]. London: British Council. Available from:https://www.britishcouncil.org/sites/default/files/linking_cse_to_uk_exams_5_0.pdf [Accessed 4th Apr 2022].

Durkin, K. (2008a) The adaptation of East Asian master's students to western norms of critical thinking and argumentation in the UK. *Intercultural Education.* **19**(1), pp.15-27.

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

Dörnyei, Z and Taguchi, T. (2010). *Questionnaires in second language research*. 2nd ed. Florence: Routledge.

Eftekhari, M., Sotoudehnama, E. and Marandi, S. S. (2016) Computer-aided argument mapping in an EFL setting: does technology precede traditional paper and pencil approach in developing critical thinking? *Educational Technology Research and Development*. **64**(2), pp.339–357.

Eisner, E. W. (1998) *The enlightened eye: qualitative inquiry and the enhancement of educational practice*. Upper Saddle River: Prentice-Hall.

Elder, L. and Paul, R. (1994) Critical thinking: why we must transform our teaching. *Journal of Developmental Education.* **18**(1), pp.34-35.

Dallimore, E. J., Hertenstein, J. H. and Platt, M. B. (2004) Classroom participation and discussion effectiveness: student-generated strategies. *Communication Education*. **53**(1), pp. 103-115.

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

Ennis, R. H. (2015) Critical thinking: a streamlined conception. In: Davis, M. and Barnett, R. (eds.) *The palgrave handbook of critical thinking in higher education*. New York: Palgrave Macmillan, pp.31-47.

Ennis, R. H. (2018) Critical thinking across the curriculum: a Vision. *Topoi*. **37**, pp.165-184.

Erstad, O. (2002) Norwegian students using digital artifacts in project-based learning. *Journal of Computer Assisted Learning.* **18**, pp. 427–437.

Eslami, Z. and Garver, R. (2013) English language learners and project-based learning. In: Capraro, R. M., Capraro, M. M. and Morgan, J. R. (eds.) *STEM project-based learning: an integrated science, technology, engineering, and mathematics.* 2nd ed. The Netherlands: Sense Publishers, pp.119-128.

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

Facione, P. A. (2015a) 2015 update on the critical thinking mindset from delphi report principle. *Insight Assessment* [online]. Available at: www.insightassessment.com [Accessed 4th January 2019].

Facione, P. A. (2015b) Critical thinking: what it is and why it counts. *Insight Assessment* [online]. Available from: www.insightassessment.com [Accessed 4th January 2019].

Facione, P. A. and Facione, N. C. (2014) Holistic Critical Thinking Scoring Rubric (HCTSR). *Insight Assessment* [online]. Available from: https://www.insightassessment.com/article/holistic-critical-thinking-scoring-rubric-hctsr [Accessed 22th October 2019].

Facione, P. A., Gittens, C. A. and Facione, N. C. (2016) Cultivating a critical thinking mindset. *Insight Assessment* [online]. Available from: http://go.roguecc.edu/sites/go.rhoguecc.edu/files/users/MWeast/Cultivating+A+Positive+Critical+Thinking+Mindset_0.pdf [Accessed 20 Feb 2022]

Fan, J. and Frost, K. (2022) At the intersection of language testing policy, practice, and research: an interview with Yan Jin. *Language Assessment Quarterly*. **19**(1), pp.76-89.

Fang, F. (2018) Ideology and identity debate of English in China: past, present and future. *Asian Englishes*. **20**(1), pp.15-26.

Farahian, M., Avarzamani, F. and Rajabi, Y. (2021) Reflective thinking in an EFL writing course: to what level do portfolios improve reflection in writing? *Thinking Skills and Creativity*. **39**, pp.1-12.

Fisher, A. (2011) *Critical thinking: an introduction*. 2nd ed. Cambridge: Cambridge University Press.

Flick, U. (2017) Triangulation in data collection. In: Flick, U. (ed.) *The SAGE handbook of qualitative data collection*. London: SAGE Publications, pp.527-544.

Flores, K. L., Matkin, G. S., Burbach, M. E., Quinn, C. E. and Harding, H. (2012) Deficient critical thinking skills among college graduates: implications for leadership. *Educational Philosophy and Theory*. **44**(2), pp.212–230.

Floyd, C. B. (2011) Critical thinking in a second language. *Higher Education Research and Development*. **30**(3), pp.289–302.

Flyvbejerg, B. (2006) Five misunderstandings about case-study research. *Qualitative Inquiry*. **12**(2), pp. 219-245.

Fong, C. J., Kim, Y., Davis, C. W., Hoang, T. and Kim, Y. W. (2017) A meta-analysis on critical thinking and community college student achievement. *Thinking Skills and Creativity*. **26**, pp.71–83.

Foss, P., Carney, N., McDonald, K. and Rooks, M. (2008) Project-based learning activities for short-term intensive English programs. In: Robertson, P. and Adamson, J. (eds.) *The Philippine ESL journal: volume 1*. Pusan: Asian EFL Journal Press, pp.57-76.

Fu, H. J. and Wang, J. C. (2021) Cultivation of critical thinking skills in college English writing under blended learning model. *Creative Education*. **12**, pp.1485-1493.

Gao, X. and Wen, Q. F. (2017) A study of the influence of critical thinking and language factors on L2 writing. *Foreign Language Teaching Theory and Practice.* **4**, pp.44-50.

Gao, Y. (2010) The application of project-based learning in college English teaching. *Foreign Language World*. **6**, pp. 42-56.

Gao, Y. H. (1999) A "short board" for foreign language learning. *Foreign Language Learning Theory and Practice*. **3**, pp.6-9.

General Office of the CPC Central Committee and General Office of the State Council. (2021) Opinions on further reducing the burden on students in the compulsory education state from homework and extracurricular training [online]. Available from: http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202107/t20210724_546576.html [Accessed 2nd November 2021].

Gibbes, M. and Carson, L. (2014) Project-based language learning: an activity theory analysis. *Innovation in Language Learning and Teaching*. **8**(2), pp.171-189.

Glasersfeld, E. V. (1998) Why constructivism must be radical. In: Larochelle, M., Bednarz, N. and Garrison, J. (eds.) *Constructivism and Education.* New York: Cambridge University Press, pp.23-28.

Gratchev, I. and Jeng, D. S. (2018) Introducing a project-based assignment in a traditionally taught engineering course. *European Journal of Engineering Education*. **43**(5), pp.788-799.

Grigg, R. and Lewis, H. (2019) *Teaching creative and critical thinking in schools*. Los Angeles: SAGE.

Gu, P. Y. (2007) Theory and practice of project-based learning with multimedia. *Foreign Language World*. **2**, pp.2-8.

Gu, P. Y. and Zhu, M. H. (2002) Online English writing and project-based learning: a case study. *Technology Enhanced Foreign Language Education*. **12**,pp.3-7.

Gu, Q. and Schweisfurth, M. (2006) Who adapts? Beyond cultural models of 'the' Chinese learner. *Language, Culture and Curriculum*. **19**(1), pp.74–89.

Haas, P. F. and Keeley, S. M. (1998) Coping with faculty resistance to teaching critical thinking. *College Teaching*. **46**, pp.63–67.

Hancock, D. R. and Algozzine, B. (2006) *Doing case study research: a practical guide for beginning researchers*. New York: Teachers College Press.

Haines, S. (1989) *Projects for the EFL classroom: resource material for teachers*. Hong Kong: Nelson.

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

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

Hamilton, L. and Corbett-Whittier, C. (2013) *Using case study in education research*. London: SAGE Publications Ltd.

Hammersley, M. (2013) *What is qualitative research?* London: Bloomsbury Academic.

Handhika, J., Cari, C., Sunarno, W., Suparmi, A. and Kurniadi, E. (2018) The influence of project-based learning on the student conception about kinematics and critical thinking skills. *Journal of Physics: Conference Series*. **1013**, pp.1-8.

Hancock, D. R. and Algozzine, B. (2006) *Doing case study research: a practical guide for beginning researchers*. New York: Teachers College Press.

Han, Y. and Hyland, F. (2015) Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing*, **30**(4), pp.31–44.

Hanjani, A. M. and Li, L. (2014) Exploring L2 writers' collaborative revision interactions and their writing performance. *System*, **44**, pp.101–114.

Harits, H. (2011) Critical thinking skills and meaning in English language teaching. *TEFLIN Journal*. **22**(2), pp.185–200.

Harizaj, M. and Hajrulla, V. (2017) Fostering learner's critical thinking skills in EFL: some practical activities. *European Scientific Journal*, **13**(29), p.126.

He, L. Z. and Lin. X. (2017) *Development of Chinese students' critical thinking through learning community in College English classroom*. Hangzhou: Zhejiang University Press.

Heard, J., Scoular, C., Duckworth, D., Ramalingam, D. and Teo, I. (2020) Critical thinking: skill development framework. *Australian Council for Educational Research.* Available from:https://research.acer.edu.au/ar_misc/41 [Accessed 18th September 2021].

Higgins, S. (2015) A recent history of teaching thinking. In: Wegerif, R., Li, L. and Kaufman, J. C. (eds.) *The Routledge international handbook of research on teaching thinking*. Abingdon: Routledge, pp.19–28.

Holt, M. (1994) Dewey and the "cult of efficiency": competing ideologies in collaborative pedagogies of the 1920s. *Journal of Advanced Composition*, **14**(1), pp.73-92.

Hou, L. Y. (2012) *Research on the teaching of English reading and the cultivation of critical thinking*. Beijing: Foreign Language Education and Research Press.

Hu, M. (2017) Teachers' intervention in developing English majors' critical thinking when teaching writing in Chinese universities. *Theory and Practice in Language Studies*. **7**(12), pp.1290–1294.

Huang, Y. S. (1998) The absence of critical thinking. *Foreign Languages and Their Teaching*. **7**, pp.18-19.

Huang, L., Fan, A. P. C., Su, N., Jessica, T., Olive, K. R. and Zhao, X. D. (2021) Chinese medical students' disposition for critical thinking: a mixed methods exploration. *BMC Medical Education*. **21**(1), pp.1–385.

Huber, C. R. and Kuncel, N. R. (2016) Does college teach critical thinking? A meta-analysis. *Review of Educational Research*. **86**(2), pp.431–468.

Hughes, J. N. and Chen, Q. (2011) Reciprocal effects of student-teacher and student-peer relatedness: effects on academic self-efficacy. *Journal of Applied Developing Psychology*, **32**(5), pp.278-287.

Hyerle, D. and Alper, L. (2011) *Student success with thinking maps.* Thousand Oaks: Corwin Press.

Ikuenobe, P. (2002) Epistemic foundation for teaching critical thinking in group discussion. *Interchange*, **33**, pp. 371-393.

Insight Assessment (2014) California critical thinking disposition inventory (CCTDI). *Insight Assessment* [online]. Available from: https://www.insightassessment.com/product/cctdi [Accessed 21th May 2019].

Insight Assessment (2020a) *California critical thinking skills test 2020 user manual and resource guide*. San Jose: The California Academic Press.

Insight Assessment (2020b) Information for understanding CCTST score report (34-point versions). Insight Assessment [online]. Available from: www.insightassessment.com [Accessed 14th April 2021].

Insight Assessment (2021) California critical thinking disposition inventory: 2021 user manual and resource guide. *Insight Assessment* [online]. Available from: www.insightassessment.com [Accessed 20nd July 2021].

Ip, W. Y., Lee, D. T. F., Lee, I. F. K., Chau, Janita, P. C., Wootton, Yvonne, S. Y. and Chang, A. M. (2000) Disposition towards critical thinking: a study of Chinese undergraduate nursing students. *Journal of Advanced Nursing*. **32**(1), pp.84–90.

Irshad, K. S., (2017) An investigation of the concept of critical thinking in the context of a functional English course in a BE degree in Pakistan. Ph.D. University of Glasgow.

Jacobs, G. M. and Farrell T. S. C. (2003) Understanding and implementing the CLT (communicative language teaching) paradigm. *RELC Journal*.. **34**(1), pp.5–30.

Jin, L. X. and Cortazzi, M. (2015) Principles, designs and practices in language teaching for developing critical thinking. In: *A Short Course on Cultivation Critical*

Thinking Skills in Foreign Language Teaching, Beijing, 2015. Beijing: Higher English Education Publishing Branch, FLTRP.

Jiang, J. Y., Quyang, J. H. and Liu, H. T. (2016) Can learning a foreign language foster analytic thinking? -Evidence from Chinese EFL learners' writings. *PLoS ONE,* **11**(10), pp.1-17.

Jones, A. (2005) Culture and context: critical thinking and student learning in introductory macroeconomics. *Studies in Higher Education*, **30**(3), pp.339–354.

Kaeppel, K. (2021) The influence of collaborative argument mapping on college students' critical thinking about contentious arguments. *Thinking Skills and Creativity*. **40**, pp.1-9.

Karbalaei, A. (2012) Critical thinking and academic achievement. *Íkala; Medellín.* **17**(2), pp. 121-128.

Kavlu, A. (2017) Implementation of Project Based Learning (PBL) in EFL (English as a Foreign Language) classrooms in Fezalar Educational Institutions (Iraq). *International Journal of Social Sciences and Educational Studies.* **4**(2), pp.67-79.

Kawson, R. (Trans.) (2000) *The Analects.* Oxford: Oxford University Press.

Khodadady, E. and Ghanizadeh, A. (2011) The impact of concept mapping on EFL learners' critical thinking ability. *English Language Teaching*. **4**(4), pp.49–60.

Khohsima, H., Zare-Behtash, E. and Khosravani, M. (2016) Examining the impact of novel pre-activity tasks on macroskills: the case of group discussion on writing ability. *Advances in Language and Literary Studies*. **7**(1), pp.148–156.

Kilpatrick, T. H. (1918) The project method. *Teachers College Record*. **19**, pp.319–34.

Kim, H. K. (2003) Critical thinking, learning and confucius: a positive assessment. *Journal of Philosophy of Education*. **37**(1), pp.71–87.

Kirby, J. R., Woodhouse, R. A. and Ma, Y. (1996) Studying in a second language: the experiences of Chinese students in Canada. In: Watkins, D. A. and Biggs, J. B. (eds.)

The Chinese learner: cultural, psychological and contextual influences. Hong Kong: Comparative Educational Research Centre, pp.141–158.

Koda, K. (2005) *Insights into second language reading.* Cambridge: Cambridge University Press.

Kokotsaki, D., Menzies, V. and Wiggins, A. (2016) Project-based learning: a review of the literature. *Improving Schools*, **19**(3), pp.267–277.

Kölbel, J. and Jentges, E. (2018) The six-sentence argument: training critical thinking skills using peer review. *Management Teaching Review*, **3**(2), pp.118–128.

Kow, H. M., (2016). Infusion of critical thinking across the English language curriculum: a multiple case study of primary school in-service expert teachers in Singapore. MEd. The University of Western Australia.

Kozulin (2018) Conceptual analysis and practice applications. In: Lantolf, J. P., Poehner, M. E. and Swain, M. (eds.) *The routledge handbook of sociocultural theory and second langugae development*. Abingdon: Routledge, pp.23-41.

Krajcik, J. S. and Shin, N. (2022) Porject-based learning. In: Sawyer, R. K. (ed.) *The cambridge handbook of the learning sciences*. 3rd ed. New York: Cambridge University Press, pp.72-92.

Ku, K. Y. L. and Ho. I. T. (2010) Dispositional factors predicting Chinese students' critical thinking performance. *Personality and Individual Differences,* **48**, pp.54–58.

Kuhd, D. (1991) *The Skills of Argument*. Cambridge: Cambridge University Press.

Kwan, Y. W. and Wong, A. F. L (2015) Effects of the constructivist learning environment on students' critical thinking ability: cognitive and motivational variables as mediators. *International Journal of Educational Research*. **70**, pp.68–79.

Lai, L. T. and Guo, F. (2018) Creating lifelong learners: using project-based management to teach 21st Century skills, *Innovations in Education and Teaching International*. **55**(6), pp.746-747.

Lantolf, J. P., and Thorne, S. L. (2006) *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press

Larmer, J., Mergendoller, J. R. and Boss, S (2015). Gold standard PBL: essential project design elements. Buck Institute for Education [online]. Available from: www.bie.org [Accessed 16 April 2022].

Lather, P. (2004) Critical inquiry in qualitative research: feminist and poststructural perspectives- science 'after truth'. In: de Marrais, K. and Lapan, S. D. (eds.) *Foundations for research: methods of inquiry in education and the social sciences.* Mahwah: Lawrence Erlbaum, pp.203-215.

Law, E. H.-F. and Miura, U. (2015) Transforming teaching and learning in Asia and the Pacific: case studies from seven countries. In: *The Third Regional High-Level Expert Meeting Beyond 2015: Transforming Teaching and Learning in the Asia-Pacific Region*, Bangkok, 2013. Paris: UNESCO.

Le, H., Janssen, J. and Wubbels, T. (2018) Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration, *Cambridge Journal of Education*. **48**(1), pp.103-122.

Lee, I. (2016) EFL writing in schools. In: Manchón, R. M. and Matsuda, P. K. (eds.) *Handbook of second and foreign language writing. Vol.11.* Boston: De Gruyter, pp.113–139.

Lee, S., (2014) CALL-infused project-based learning: a case study of adult ESL students learning prepositions. MA. Iowa State University.

Lee, W. W. S. (2018) The timing and critical incident of epistemic beliefs change in Hong Kong college students: an exploratory study. *Asia Pacific Journal of Education*. **38**(2), pp.164–174.

Lee, Y. (2013) Collaborative concept mapping as a pre-writing strategy for L2 learning: a Korean application. *International Journal of Information and Education Technology*. **3**(2), pp.254–258.

Leggett. G. and Harrington, I. (2021) The impact of project based learning (PBL) on students from low socio economic statuses: a review, *International Journal of Inclusive Education*. **25** (11), pp.1270-1286.

Lehmann, .M., Christensen, P., Du, X. and Thrane, M. (2008) Problem-oriented and Project-Based Learning (POPBL) as an innovative learning strategy for sustainable development in engineering education. *European Journal of Engineering Education.* **33** (3), pp.283–295.

Li, H. H. and Zhang, L. J. (2021) Effects of structured small-group student talk as collaborative prewriting discussions on Chinese university EFL students' individual writing: a quasi-experimental study. *PloS One*. **16**(5), pp.1-20.

Li, H. H., Zhang, L. J. and Parr, J. M. (2020) Small-group student talk before individual writing in tertiary English writing classrooms in China: nature and insights. *Frontiers in Psychology*. **11**, pp.1-14.

Li, J. (2017) Another reason for the "critical absence" in English writing: the perspective of language testing. *Journal of Zibo Normal College*. **4**, pp. 34-38.

Li, K. (2010) Project-based college English: an approach to teaching non-English majors. *Chinese Journal of Applied Linguistics(Biomonthly)*. **33**(4), pp.99-122.

Li, L. (2011) Obstacles and opportunities for developing thinking through interaction in language classrooms. *Thinking Skills and Creativity.* **6**(3), pp.146–158.

Li, L. (2016a) Thinking skills and creativity in second language education: where are we now? *Thinking Skills and Creativity*. **22**, pp.267–272.

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

Li, L. (2020) "Thinking" move in second language education. In: Li, L. (ed.) *Thinking skills and creativity in second language education*. Abingdon: Routledge, pp.1-16.

Li, L. and Du, J. (2014) PBL mode of academic English in discipline-based college English teaching. *Foreign Language Education*. **35**, pp.55-58.

Li, L. and Wegerif, R. (2014) What does it mean to teach thinking in China? Challenging and developing notions of 'Confucian education'. *Thinking Skills and Creativity*. **11**, pp.22–32.

Li, L. W. (2011) An action research on how to increase reader awareness and critical thinking. *Foreign Language in China*. **8**(3), pp.66-73.

Li, M., Zhao, G. F. and Cao, G. G. (2016) A survey report of Chinese College English course. *Theory and Practice in Language Studies*. **6**(8), pp.1557-1565.

Li, X. L. and Liu, J. (2021) Mapping the taxonomy of critical thinking ability in EFL. *Thinking Skills and Creativity*. **41**, pp.1-13.

Liang, W. and Fung, D. (2021) Fostering critical thinking in English-as-a-second-language classrooms: challenges and opportunities. *Thinking Skills and Creativity*. **39**, pp.1-12.

Liberante, L. (2012) The importance of teacher-student relationships, as explored through the lens of The NSW quality teaching model. *Journal of Student Engagement: Education Matters*. **2**(1), pp.2-9.

Lin, M. and Xiang, X. (2020) Integrating critical thinking into an EFL writing curriculum: a mediated model. In: Li L. (ed.) *Thinking skills and creativity in second language education: case studies from international perspectives*. Abingdon: Routledge, pp.95-130.

Lin, Y. (2012) Oral teaching and critical thinking training. *Foreign Language and Their Teaching.* **5**, pp.29-33.

Lin, Y., (2014) Infusion of critical thinking into L2 classes: a case study in a Chinese high school. Ph.D. Newcastle University.

Lin, Y. (2018) *Developing critical thinking in EFL classes: an infusion approach*. Singapore: Springer.

Lincoln, Y. S. and Guba, E. G. (1985) *Naturalistic Inquiry*. Newbury Park: Sage.

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

Liu, P. (2011) A study on the use of computerized concept mapping to assist ESL learners' writing. *Computers and Education*. **57**(4), pp.2548-2558.

Liu, H. and Jin, L. M. (2012) An empirical study of English and the development of critical thinking in college students. *Foreign Language and Their Teaching*. 5, pp.24-28.

Liu, J. (2018) Cultivation of critical thinking abilities in English writing teaching. *Theory and Practice in Language Studies*. **8**(8), pp.982–987.

Lucas, K. J. (2019) Chinese graduate student understandings and struggles with critical thinking: a narrative-case study. *International Journal for the Scholarship of Teaching and Learning*. **13**(1), pp.1-7.

Luk, J. (2012) Teachers' ambivalence in integrating culture with EFL teaching in Hong Kong. *Language, Culture and Curriculum*. **25**, pp.249–264.

Luk, J. and Lin, A. (2015) Voices without words: doing critical literate talk in English as a second language. *TESOL Quarterly*. **49**(1), pp.67–91.

Lukenchuk, A. (2013) *Paradigms of Research for the 21st Century*. New York: Peter Lang.

Lun, V. M. C., Fischer, R., and Ward, C. (2010) Exploring cultural differences in critical thinking: is it about my thinking style or the language I speak? *Learning and Individual Differences*. **20**, pp. 604-616.

Luo, Q. X. and Yang, X. H. (2001) Revising on Chinese version of California Critical Thinking Disposition Inventory. *Psychological Development and Education*. **17**(3), pp.47-51.

Luo, Q. X. and Yang, X. H. (2002) Revising on Chinese version of California Critical Thinking Skills Test. *Psychological Science*. **25**(6), pp.740-741.

Ma, L. and Luo, H. (2021) Chinese pre-service teachers' cognitions about cultivating critical thinking in teaching English as a foreign language. *Asia Pacific Journal of Education*. **41**(3), pp.543–557.

Mackey, A. and Gass, S. M. (2005) *Second Language Research: Methodology and Design*. New Jersey: Routledge Taylor and Francis Group.

Manalo, E. and Sheppard, C. (2016) How might language affect critical thinking performance? *Thinking Skills and Creativity*. **21**, pp.41–49.

Manalo, E., Kusumi, T., Koyasu, M., Michita, Y. and Tanka, Y. (2015) Do students from different cultures think differently about critical and other thinking skills? In: Davies, M. and Barnett, R. (eds). *The Palgrave handbook of critical thinking in higher education.* Basingstoke: Palgrave Macmillan, pp. 299-316.

Magnusson, E. and Marecek, J. (2015) *Doing interview-based qualitative research*: *a learner's guide*. Cambridge: Cambridge University Press.

Markham, T., Larmer, J. and Ravitz, J. (2003) *Project based learning handbook: a guide to standards-focused project based learning for middle and high School teachers*. Hong Kong: Buck Institute for Education.

Martinez, C. (2022) Developing 21st century teaching skills: a case study of teaching and learning through project-based curriculum, *Cogent Education*. **9**(1), pp.1-16.

Marzban, A. and Barati, Z. (2016) On the relationship between critical thinking ability, language learning strategies, and reading comprehension of male and female intermediate EFL university students. *Theory and Practice in Language Studies*. **6**(6), pp.1241–1247.

Maxwell, J. A. (2005) *Qualitative research design: an Interactive Approach*. 2nd ed. Thousand Oaks: Sage.

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

McGuire, J. M. (2007) Why has the critical thinking movement not come to Korea? *Asia Pacific Education Review*. **8**(2), pp.224–232.

McKinley, J. (2013) Displaying critical thinking in EFL academic writing: a discussion of Japanese to English contrastive rhetoric. *RELC Journal*. **44**(2), pp.195–208.

McPeck, J. E. (1981) *Critical thinking and education*. New York: St Martin's Press.

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

McMillan, J. H. and Schumacher, S. (2010) *Research in education: evidence-based inquiry*. 7th ed. New Jersey: Pearson Education.

Megayanti, T., Busono, T. and Maknun, J. (2020) Project-based learning efficacy in vocational education: literature review. *Materials Science and Engineering*. **830**(4), pp.1-6.

Melles, G. (2009) Teaching and evaluation of critical appraisal skills to postgraduate ESL engineering students. *Innovations in Education and Teaching International*. **46**(2), pp.161–170.

Mercer, N., Hennessy, S. and Warwick, P. (2019) Dialogue, thinking together and digital technology in the classroom: some educational implications of a continuing line of inquiry. *International Journal of Educational Research*. 97, pp.187-199.

Mergendoller, J. (2012) Does project based learning teach critical thinking? *Buck Insttute for Educaton* [online]. Available from: http://tc2.ca/uploads/PDFs/Critical%20Discussions/does_project_based_learning_teach_critical_thinking.pdf [Accessed 11 June 2019]

Merriam, S. B. and Tisdell, E. J. (2015) *Qualitative research: a guide to design and implementation*. 4th ed. San Francisco: Jossey-Bass.

Meyer, E. and Forester, L. A. (2015) Implementing student-produced video projects in language courses. *Unterrichtspraxis*. **48**(2), pp.192–210.

Minakova, L. Y. (2014) Critical thinking development in foreign language teaching for non-language-majoring students. *Procedia- Social and Behavioral Sciences.* **154**(C), pp.324–328.

MOE of the People's Republic of China (2007) *College English Curriculum Requirements*. (Chairman: Sir Ji Zhou.) Shanghai: Shanghai Foreign Language Education Press.

MOE of the People's Republic of China (2018) An implementation notice on vigorously grasping the spirit of national undergraduate education work conference in the new era. *MOE* [online]. Available from: http://www.moe.gov.cn/srcsite/A08/s7056/201809/t20180903_347079.html [Accessed 04 May 2022].

Moon, J. (2008) *Critical thinking: an exploration of theory and practice*. New York: Routledge.

Moore, T. J. (2004) The critical thinking debate: how general are general thinking skills? *Higher Education: Research and Development*. **23**(1), pp.3–18.

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

Moore, T. J. (2013) Critical thinking: seven definitions in search of a concept. *Studies in Higher Education*. **38**(4), pp.506–522.

Moore, T. J. (2015) Knowledge, disciplinarity and the teaching of critical thinking. In: Wegerif, R., Li, L. and Kaufman, J. C. (eds.) *The Routledge international handbook of research on teaching thinking*. Abingdon: Routledge, pp. 243–253.

Moritoshi, T. P. (2017) Project-based language learning: a case study at a Japanese junior college. Ph.D. Aston University.

Morse, D. and Jutras, F. (2008) Implementing concept-based learning in a large undergraduate classroom. *CBE Life Sci Educ*. **7**(2), pp.243-253.

Mukama, E. (2014) Bringing technology to students' proximity: a sociocultural account of technology-based learning projects. *International Journal for Research in Vocational Education and Training (IJRVET)*. **1**(2), pp.125-141.

Mulnix, J. W. (2012) Thinking critically about critical thinking. *Educational Philosophy and Theory*. **44**(5), pp.464–479.

National Bureau of Statistics of China (2002) Internet Penetration Rate. *National Bureau of Statistics of China* [online]. Available from: http://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0G0X&sj=2000 [Accessed 10th June 2019].

National Language Commission of the People's Republic of China (2018) *China's Standards of English Language Ability*. (Chairperson: Sir Xuejun Tian.) Beijing: Higher Education Press.

Ng, C. H., Cheung, Y. L., Zhang, W. Y. and Jang, H. (2021) Researching project-based learning: a review and a look ahead at form and function in writing. In: Beckett, G. H. and Slater, T. (eds.) *Global perspectives on project-based language learning, teaching, and assessment: key approaches, technology tools, and frameworks*. Abingdon: Routledge, pp.23-46.

Nagjie, B. N., Li, Y., Tiruneh, D. T. and Cheng, M. (2020) Investigating the effects of a systematic and model-based design of computer-supported argument visualization on critical thinking. *Thinking Skills and Creativity*. **38**, pp. 1-15.

Niu, L., Behar-Horenstein, L. S. and Garvan, C. W. (2013) Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review*. **9**, pp.114–128.

Noone, T. and Seery, A. (2018) Critical thinking dispositions in undergraduate nursing students: a case study approach. *Nurse Education Today*. **68**, pp.203–207.

Norris, S. P. (1985) Synthesis of research on critical thinking. *Educational Leadership*. **42**(8), pp.40-45.

Nosich, G. M. (2012) *Learning to think things through: a guide to critical thinking across the curriculum*. 4th ed. Boston: Pearson.

Nussbaum, E. M. (2002) The process of becoming a participant in small-group critical discussions: a case study. *Journal of Adolescent and Adult Literacy*. **45**(6), pp.488–97.

Oatley, K. and Djikic, M. (2008) Writing as thinking. *Review of General Psychology*. **12**(1), pp.9–27.

Okolie, U. C., Oluka, B. N., Oluwayemisi, F. B., Achilike, B. A. and Ezemoyih, C. M. (2022) Overcoming obstacles to collaborative learning practices: a study of student learning in higher education-based vocational education and training. *Internal Journal of Training Research*. **20**(1), pp.73-91.

Ozogul, G. and Sullivan, H. (2007) Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*. **57**(3), pp.393-410.

Pallant, J. (2016) *SPSS survival manual: a step by step guide to data analysis using IBM SPSS*. 6th ed. Berkshire: Open University Press.

Pallant, J. (2020) *SPSS survival manual: a step by step guide to data analysis using IBM SPSS*. 7th ed. London: Open University Press.

Pally, M. (2001) Skills development in 'sustained' content-based curricula: case studies in analytical/critical thinking and academic writing. *Language and Education*. **15**(4), pp.279–305.

Panhwar, A. H., Ansari, S. and Ansari, K. (2016) Sociocultural theory and its role in the development of language pedagogy. *Advances in Language and Literary Studies*. **7**(6), pp.183-188.

Papanikolaou, K. A. and Gouli, E. (2013) Investigating influences among individuals and groups in a collaborative learning setting. *International Journal of E-Collaboration*. **9**(1), pp.9-25.

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

Paul, R. W., Elder, L. and Bartell, T. (1997) California teacher preparation for instruction in critical thinking: research findings and policy recommendations. *California Commission on Teacher Credentialing* [online]. Available from: https://files.eric.ed.gov/fulltext/ED437379.pdf [Accessed 29th April 2019].

Paton, M. (2005). Is critical analysis foreign to Chinese students? *Google Scholar* [online]. Available from: https://s3.amazonaws.com/academia.edu.documents/31449442/PatonTonga.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1551279660&Signature=MrZDYs%2FWWORYN86Y%2BTy5%2Bq5kbUM%3D&response-content-disposition=inline%3B%20filename%3DIs_Critical_Analysis_Foreign_to_Chinese.pdf [Accessed 27 February 2019].

Pei, Z. W., Zheng, C. Q., Zhang M. and Liu F. Z. (2017) Critical thinking and argumentative writing: inspecting the association among EFL learners in China. *English Language Teaching*. **10**(10), pp.31-42.

Pellegrino, J. W. and Hilton, M. L. (2012) *Education for life and work: developing transferable knowledge and skills in the 21st century.* Washington DC: National Academies Press.

Pishghadam, R. and Ghanizadeh, A. (2006) On the impact of concept mapping as a prewriting activity on EFL learning's writing ability. *Iranian Journal of Applied Linguistics*. **9**(2), pp.101-126.

Pithers, R. T. and Soden, R. (2000) Critical thinking in education: a review. *Educational Research*. **42**, pp.237–249.

Poole, D. (2000) Student participation in a discussion-oriented online course: a case study. *Journal of Research on Computing in Education*. **33**(2), pp.162-177.

Praba, L. T., Artini, L. P. and Ramendra, D. P. (2017) Project-based learning and writing skill in EFL: are they related? *SHS Web of Conferences*. **42**, pp.1-6. .

Qian, D. D. and Cumming, A. (2017) Researching English language assessment in China: focusing on high-stakes testing. *Language Assessment Quarterly*. **14**(2), pp.97–100.

Qin, X. Q. and Bi, J. (2015) *Quantitative approaches and quantitative data analysis in L2 research.* Beijing: Foreign Language Teaching and Research Press.

Rafi, M. S. (2011) Promoting critical pedagogy in language education. *International Research Journal of Arts and Humanities*. **39**, p.105-116.

Rahman, R. and Ambreen, M. (2018) Concept mapping for improving expository writing in second language. *Pakistan Journal of Education*. **35**(2), pp.17-36.

Rastall, P. (2006) Introduction: the Chinese learner in higher education- transition and quality issues. *Language*, *Culture and Curriculum*. **19**(1), pp.1-4.

Ravitch, D. (2000) *Left back: A century of failed school reforms*. New York: Simon and Schuster.

Ren, W. (2013) Revisiting foreign language majors' critical thinking skills- a case study of English public speaking course. *Foreign Language in China*. **10**(1), pp.10-17.

Research Ethics Committee of the University of Northampton (2022) *Research Ethics Code and Procedures Version 3.0*. Available from: https://www.northampton.ac.uk/wp-content/uploads/2021/04/uon-research-ethics-code-and-procedures-may-2022.pdf [Accessed 06 July 2022].

Rezaei, S., Derakhshan, A. and Bagherkazemi, M. (2011) Critical thinking in language education. *Journal of Language Teaching and Research.* **2**(4), pp.769–777.

Robinson, P. (2013) *The routledge encyclopedia of second language acquisition.* London: Routledge.

Rochmahwati, P. (2015) Fostering students' critical thinking by project-based learning. *Journal on English as a Foreign Language*. **5**(1), pp.37–44.

Rosen, Y. and Tager, M. (2014) Making student thinking visible through a concept map in computer-ased assessment of critical thinking. *Journal of Educational Computing Research*. **50**(2), pp.249–270.

Rouvrais, S., Ormrod, J., Landrac, G., Mallet, J., Gilliot, J. M., Thepaut, A. and Tremenbert, P. (2006) A mixed project-based learning framework: preparing and developing student competencies in a French "Grande Ecole". *European Journal of Engineering Education*. **31**(1), pp.83–93.

Saad, A. and Zainudin, S. (2022) A review of Project-Based Learning (PBL) and Computational Thinking (CT) in teaching and learning. *Learning and motivation*. **78**, pp.1-12.

Sadeghi, B., Hassani, M. T. and Rahmatkhah, M. (2014) The relationship between EFL learners' metacognitive strategies, and their critical thinking. *Journal of Language Teaching and Research*. **5**(5), pp.1167–1175.

Saenabl, S., Yunus, S. R., Saleh, A. R., Virninda, A. N., L, H. and Sofyan, N. (2018) Project-based learning as the atmosphere for promoting students' communication skills. *Journal of Physics: Conference Series*. **1028**(1), pp.1-6.

Saldaña, J. (2011) *Fundamentals of qualitative research.* London: SAGE Publications Ltd..

Saldaña, J. (2016) *The coding manual for qualitative researchers.* New York: Oxford University Press.

Sasson, I., Yehuda, I. and Malkinson, N. (2018) Fostering the skills of critical thinking and question-posing in a project-based learning environment. *Thinking Skills and Creativity*. **29**, pp.203–212.

Senthamarai, T. and Chandran, M. R. (2016) Integrating critical thinking in language learning through project based learning. *International Journal of Arts, Humanities and Management Studies*. **2**(2), pp.61-65.

Shi, X. (2009) A study on project-based college English autonomous learning. *Foreign Language Education in China.* **2**, pp.16-27.

Shirkhani, S. and Fahim, M. (2011) Enhancing critical thinking in foreign language learners. *Procedia - Social and Behavioral Sciences*. **29**, pp.111-115.

Sidman-Taveau, R. L., (2005). Computer-assisted project-based learning in second language: case studies in adult ESL. Ph.D. University of Texas at Austin.

Silva, H., Lopes, J., Dominguez, C., Payan-Carreira, R., Morais, E., Nascimento, M. and Morais, F. (2016) Fostering critical thinking through peer review between cooperative learning groups. *Revista Lusófona De Educação*. **32**, pp.31-45.

Simpson, J., (2011) Integrating project-based learning in an English tourism classroom in a Thai university. Ph.D. Australian Catholic University.

Slater, T. and Beckett, G. H. (2019) Integrating language, content, technology, and skills development through project-based language learning: blending frameworks for successful unit planning. *Mextesol Journal*. **43**(1), pp.1-14.

Snider, D. (2017) Critical thinking in the foreign language and culture curriculum. *The Journal of General Education.* **66**(1-2), pp.1–16.

Sofo, F. (2004) *Open Your Mind: The 7 Keys to Thinking Critically.* Crows Nest: Allen and Unwin.

Song, X. (2016) Critical thinking and pedagogical implications for higher education. *East Asia*. **33**(1), pp.25–40.

Soufi, N. E. and See, B. H. (2019) Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Studies in Educational Evaluation*. **60**, pp.140–162.

Spencer-Oatey, H. and Xiong, Z. (2006) Chinese students' psychological and sociocultural adjustments to Britain: an empirical study. *Language, Culture and Curriculum*. **19**(1), pp.37–53.

Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students. *Written Communication*. **18**(4), pp.506-548.

Stapleton, P. (2010) A survey of attitudes towards critical thinking among Hong Kong secondary school teachers: implications for policy change. *Thinking Skills and Creativity*. **6**, pp.14-23.

Stoller, L. F. (2002) Project work: a means to promote language and content. In: Richards, J. C. and Renandya, W. A. (eds.) *Methodology in language teaching: an anthology of current practice.* New York: Cambridge University Press, pp.107-119.

Stoller, L. F. (2006) Establishing a theoretical foundation for project-based learning in second and foreign language contexts. In: Beckett, G. H. and Miller, P. C. (eds.) *Project-based second and foreign language education: past, present and future*. Greenwich: Information Age Publishing Inc, pp.19-40.

Storch, N. (2017) Sociocultural theory in the L2 classroom. In: Loewen, S. and Sato, M. (eds.) *The Routledge Handbook of Instructed Second Language Acquisition*. Abingdon: Rouledge, pp.69-84.

Storch, N. and Wigglesworth, G. (2018) Writing tasks: the effects of collaboration. In: García Mayo, M. P. (ed.) *Investigating tasks in formal language learning*. Clevedon: Cromwell Press Ltd, pp.157–177.

Sullivan, G. M. and Feinn, R. (2012) Using effect size- or why the p value is not enough. *Journal of Graduate Medical Education*. **4**(3), pp.279–282.

Sun, W. (2017) *A case study on the development of critical thinking ability of Chinese college English speech learners.* Beijing: Foreign Language Teaching and Research Press.

Sun, Y. Z. (2011) Discussion on the teaching of English writing and the cultivation of critical thinking. *Foreign Language Teaching and Research*. **43**(3), pp.603-608.

Sun, Y. Z. (2015) Foreign language education and the cultivation of critical thinking ability. *Foreign Languages in China*. **12**(2), pp.22-23.

Sun, Y. (2019) Principles of language and critical thinking integrated teaching: TERRIFIC. *Foreign Language Teaching and Research*. **51**(6), pp.825–837.

Swain, M., Kinnear, P. and Steinman, L. (2015) *Sociocultural theory in second language education: an introduction through narratives*. 2nd ed. Bristol: Multilingual Matters.

Tabak, I. and Reiser, B. J. (2022) Scaffolding. In: Sawyer, R. K. (ed.) *The cambridge handbook of the learning sciences*. 3rd ed. New York: Cambridge University Press, pp.53-71.

Tan, C. (2017) A Confucian conception of critical thinking. *Journal of Philosophy of Education*. **51**(1), pp.331–343.

Tanaka, J. and Gilliland, B. (2017) Critical thinking instruction in English for academic purposes writing courses: a dialectical thinking approach. *TESOL Journal*. **8**(3), pp.657–674.

Tang, C. and Biggs, J. (1996) How Hong Kong students cope with assessment. In: Watkins, D. A. andBiggs, J. B. (eds.) *The Chinese learner: cultural, psychological, and contextual influences*. Hong Kong/Melbourne: CERC/ACER, pp.159-182.

Tavakol, M. (2018) Coefficient alpha. In: Frey, B. (ed.) *The SAGE encyclopedia of educational research, measurement, and evaluation*. Thousand Oaks: Sage, pp.303-306.

Teng, L. S. and Zhang, L. J. (2017) Effects of motivational regulation strategies on writing performance: a mediation model of self-regulated learning of writing in

English as a second/foreign language. *Metacognition and Learning*. **13**(2), pp.213-240.

The National Foreign Languages Teaching Advisory Board (2016) *College English curriculum requirements*. (Chairperson: Madam Lianzhen He.) Shanghai: Shanghai Foreign Language Education Press.

The National Foreign Language Teaching Advisory Board (2020) *College English curriculum requirements*. (Chairperson: Madam Lianzhen He) Shanghai: Shanghai Foreign Language Education Press.

Thomas, E. and Magilvy, J. K. (2011) Qualitative rigor or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing*. **16**(2), pp.151–155.

Thomas, J. W. (2000) A review of research on project-based learning. *Google Scholar* [online]. Available from: https://www.asec.purdue.edu/lct/HBCU/documents/AReviewofResearchofProject-BasedLearning.pdf [Accessed 2 June 2019].

Tian, J., (2008) The influence of undergraduate learning contexts on Chinese graduate students' argumentation and critical thinking in writing. Ph.D. University of York.

Tian, J. and Low, G. D. (2011) Critical thinking and Chinese university students: a review of the evidence. *Language, Culture and Curriculum*. **24**(1), pp.61–76.

Tight, M. (2017) Case study research. In: Wyse, D., Selwyn, N., Smith, E. and Suter, L. E. (eds.) *The BERA/SAGE handbook of educational research: two volume set.* London: SAGE Publications Ltd.

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

Tracy, S. J. (2010) Qualitative quality: eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*. **16**(10), pp.837–851.

Tricia, H. (1993) Key concepts in ELT. *ELT Journal*. **47**(3), pp.275–277.

Trilling, B. and Fadel, C. (2009) *21st century skills: learning for life in our times.* San Francisco: Jossy-Bass.

Trobia, A. (2011) Cronbach Alpha. In: Lavrakas, P. J. (ed.) *Encyclopedia of survey research methods*. Thousand Oaks: Sage, pp.169-170.

Tsui, L. (1999) Courses and instruction affecting critical thinking. *Research in Higher Education*. **40**(2), pp.185–200.

Tsui, L. (2002) Fostering critical thinking through effective pedagogy. *The Journal of Higher Education*. **73**(6), pp.740–763.

Turner, Y. (2006) Students from mainland China and critical thinking in postgraduate business and management degrees: teasing out tensions of culture, style and substance. *International Journal of Management Education*. **5**(1), pp.3–11.

van Compernolle, R. A. and Williams, L. (2013) Sociocultural theory and second language pedagogy. *Language Teaching Research*. **17**(3), pp.277-281.

Villarreal, I. and Gil-Sarratea, N. (2020) The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research: LTR*. **24**(6), pp.874–897.

Vorobel, O. and Kim, D. (2017) Adolescent ELLs' collaborative writing practices in face-to-face and online contexts: from perceptions to action. *System*. **65**, pp.78–89.

Wan, G. F. (2001) The learning experience of Chinese students in American universities: a cross-cultural perspective. *College Student Journal*. **35**(1), pp.28.

Wang, B. (2012) On students' satisfaction at learning achievement in the project-based college English learning model-a study of PBL from the perspective of social constructivism. *Journal of Northeastern University*. **14** (5), pp.461-466.

Wang, B. (2013) An analysis of the effects of PBL model on college English learning motivation. *Technology Enhanced Foreign Language Education*. **149**, pp.37-41 and pp.68.

Wang, H. (2013) An application of PBL into the Business English course. *China University Teaching*. **8**, pp.52-54.

Wang, J. Q. and Wen Q. F. (2011) Introduction of critical thinking instruments from abroad and some inspirations. *Journal of Jiangsu Teachers University of Technology*. **17**(7), pp.38-42.

Wang, S. (2017) An exploration into research on critical thinking and its cultivation: an overview. *Theory and Practice in Language Studies*. **7**(12), pp.1266–1280.

Wang, S. (2020) Project-based language learning in China: a literature review. *Journal of Language Teaching and Research*. **11**(1), pp.66-72.

Wang, S. and Seepho, S. (2017) Facilitating Chinese EFL learners' critical thinking skills: the contributions of teaching strategies. *SAGE open*, **7**(3), p.1-9.

Wang, S. J. and Zhang, Y. (2017) Study of writing problem in College General English course—reflection on the reform of College English course. *Journal of Language Teaching and Research*. **8**(1), p.176-183.

Webb, N. L. (2004) Mathematics education reform in California. In: *Science and mathematics education in the United States: eight innovations: proceedings of a conference*, Paris, 2000. Paris: OECD.

Wegerif, R., Mercer, N. and Dawes, L. (1999) From social interaction to individual reasoning: an empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction*. 9, pp. 493-516.

Wei, R. and Su, J. (2015) Surveying the English language across China. *World Englishes*. **34**(2), pp.175-189.

Wen, Q. F., Wang, H. M., Wang, J. Q., Zhao, C. R. and Liu, Y. P. (2010) A comparative study of critical thinking skills between English and other liberal arts majors. *Foreign Language Teaching and Research*. **42**(5), pp.350-355.

Wen, Q. F. and Zhou, Y. (2006) Comment on the development of thinking ability of foreign language major students. *Foreign Language Research*. **5**, pp.76-80.

Wen, Q. F., Sun, W. and Zhang, L. L. (2018) A study on the trend of the development of critical thinking skills of foreign language majors. *Foreign Language World*. **6**, pp. 12-19.

Wen, Q. F., Zhang, L. L. and Sun, W. (2014) Are foreign language majors inferior to the other majors in critical thinking? *Modern Foreign Languages*. **37**(6), pp.794-804.

Woodward-Kron, R. (2002) Critical analysis versus description? Examining the relationship in successful student writing. *Journal of English for Academic Purposes*. **1**(2), pp.121–143.

Wertsch, J. V. (2007) Mediation. In: Daniels, H., Cole, M. and Wertsch, J. V. (eds.) *The Cambridge companion to Vygotsky*. Cambridge: Cambridge University Press, pp.178-192.

Wright, S. and Zheng, L. (2016) English in Chinese higher education: past difficulties, current in intiatives and future challenges. *Journal of World Languages*. 3(3), pp.167-182.

Wu, H. (2019) Reticence in the EFL classroom: voices from students in a Chinese university. *International Journal of Applied Linguistics and English Literature*. **8**(6), pp.114–125.

Wu, Y. (2018) Constructing Chinese gold courses. 24 November. *The 11th China University Teaching Forum* [online]. Available from: http://www.sohu.com/a/277768529_667940 [Accessed from 19 March 2021]

Xia, S. H. and Zhang, W. Z. (2017) Learner effort in project-based language learning. *Foreign Language and Their Teaching*. **3**, pp.78-88.

Yang, L. and Han, G. (2012) An empirical study of the PBL approach to academic college English writing. *Foreign Language World*. **5**, pp.8-16.

Yang, Y-T. C. and Gamble (2013) Effective and practical critical thinking enhanced EFL instruction. *ELT Journal*. **67**(4), pp.398-412.

Yang, Y-T. C., Newby, T. and Bill, R. (2008) Facilitating interactions through structured web-based bulletin boards: a quasi-experimental study on promoting learners' critical thinking skills. *Computers and Education*. **50**(4), pp.1572–1585.

Yeh, M. L and Chen, H. H. (2003) Comparison affective dispositions toward critical thinking across Chinese and American baccalaureate nursing students. *Journal of Nursing Research*. **11**(1), pp. 39–46.

Yin, R. K. (2006) Mixed methods research: are the methods genuinely integrated or merely parallel? *Research in the Schools*. **13**(1), pp.41–47.

Yin, R. K. (2018) *Case study research and applications: design and methods*. 6th ed. Los Angeles: SAGE.

Yin, Y. (2018) *Developing critical thinking in EFL classes-an infusion approach*. Singapore: Springer.

Yu, Q., (2012) Research on the application of project-based instruction design in English newspaper instruction. MA. Shanghai International Studies University.

Yu, W. (2017) Students' perceptions of the impacts of project activities in learning New Experiencing English. *Foreign Languages in China.* **14** (3), pp.82-91.

Yu, S., Geng, F., Liu, C. and Zheng, Y. (2021) What works may hurt: the negative side of feedback in second language writing. *Journal of Second Language Writing*. **54**, pp.1-15.

Yuan, R., and Stapleton, P. (2020). Student teachers' perceptions of critical thinking and its teaching. *ELT Journal*, **74**(1), pp.40–48.

Zabit, M. N. M., Karagiannidou, E. and Zachariah, T. Z. (2016) Teaching business in Malaysia and the use of PBL to nurture students' critical thinking: a case study of Sultan Idris Education University. *Malaysian Journal of Society and Space*. **12**(6), pp. 1-11.

Zare, P. (2015) Critical thinking skills among EFL/ESL learners: a review of literature. *Language in India*. **15**(11), pp.241-257.

Zhai, M. (2021) Collaborative writing in a Chinese as a foreign language classroom: learners' perceptions and motivations. *Journal of Second Language Writing*. **53**, p.100836.

Zhang, H., Yuan, R. and He, X. (2020) Investigating university EFL teachers' perceptions of critical thinking and its teaching: voices from China. *The Asia-Pacific Education Researcher*. **29**(5), pp.483–493.

Zhang, L. J. (2013) Second language writing as and for second language learning. *Journal of Second Language Writing*. **22**, pp.446–447.

Zhang, T. (2016) Why do Chinese postgraduates struggle with critical thinking? Some clues from the higher education curriculum in China. *Journal of Further and Higher Education*. **41**(6), pp. 857-871.

Zhang, W. Z. (2015) iPBL: a localized project-based English teaching and learning model. *Foreign Languages in China*. **12**(2), pp.15-23.

Zhang, Y. Q., Li, L. S, Wu, P. and Chen, Y. (2009) Investigation and analysis of critical thinking ability in medical students. *Shanghai Jiaotong University (Med Sci)*. **29**(7), pp.869–872.

Zhang, Z. and Hyland, K. (2018) Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*. **36**, pp.90–102.

Zhao, J. (2017) Confucius as a critical educator: towards educational thoughts of Confucius. *Frontiers of Education in China*. **8**(1), pp.9–27.

Zhao, W. (2020) Epistemological flashpoint in China's classroom reform: (How) can a 'Confucian do-after-me pedagogy' cultivate critical thinking? *Journal of Curriculum Studies*. **52**(1), pp.101–117.

Zhao, Z. T. (2012) EFL teaching and reform in China's tertiary education. *Journal of Language Teaching and Research*. **3**(6), pp.1098-1104.

Zhong, W. C. and Cheng, M. (2021) Developing critical thinking: experiences of Chinese international students in a post-1992 university in England. *Chinese Education and Society*. **54**, pp.3-4, 95-106.

Zhou, J., Jiang, Y. H. and Yao, Y. (2015) The investigation on critical thinking ability in EFL reading class. *English Language Teaching.* **8**(1), pp.83–94.

Zhou, Y. X. (2008) A review of education in China. *Journal of Hefei University of Technology*. **22**(2), pp.166-170.

Zhou, Z. (2018) A study on the cultivation of critical thinking ability of English majors. *Theory and Practice in Language Studies*. **8**(3), pp.349-353.

Zhu, X. L. and Shen, N. (2004) Research on the reliabilities and validities of Watson-Glaser Critical Thinking Appraisal and California Critical Thinking Disposition Inventory. *Journal of Nursing Science*. **19**(21), pp. 56-58.

Zou, M. and Lee, I. (2021) Learning to teach critical thinking: testimonies of three EFL teachers in China. *Asia Pacific Journal of Education*. **9**, pp.1-15.

Appendixes

Appendix A. The two-dimensional Structure Model of Critical Thinking

(Facinoe, 1990, p.4-13; Facino, A., Gittens, Facione, N., 2016, p.8)

Core cognitive abilities ⟶	Sub-skills	Dispositions
Interpretation	Categorization: • to provide a categorical framework for understanding, describing or characterizing information.	**Truth-seeking:** • the person would like to pursue as much knowledge as possible with a strong desire and keeps honest even in the face of biases, prejudices or stereotypes. **Open-mindedness:** • the person has an open-mindedness regarding different worldviews and develops an understanding of the opinions from other people. **Analyticity:** • the person could make judgements predicatively based on their trust in the processes of reasoned inquiry. **Systematicity:** • the person is liable to identify and solve problems step by step on the basis of a methodic and radical approach. **Confident in Reasoning:** • the person has self-confidence in one's own abilities to reason in order to make a logical judgement. **Inquisitiveness:** • the person possesses inquisitiveness with regard to a wide range of issues and expresses concerns to
	Decoding Significance: • to discover and describe the content, meaning, functions, motives, values and rules etc.	
	Clarifying Meaning: • to interpret and clarify the meanings of views, concepts, behaviours and symbols, finally to remove ambiguity.	
Analysis	Examining ideas: • to define terms; • to compare or contrast ideas, concepts, or statements; • to identify the relationships between the part and the whole.	
	Detecting arguments: • to detect whether a set of statements, including assertions and corresponding justifications, could form arguments.	
	Analyzing arguments: • to divide the reasoning process into presuppositions, premises and conclusions.	
Evaluation	Assessing Claims: • to assess judgements and the reliability, relevance and credibility to sources of information.	
	Assessing Arguments: • to judge whether presuppositions and premises of a given argument are true or not, whether arguments contain fallacy and to judge strength of arguments.	

Inference	Querying Evidence: • to identify, search and filter the evidence to support arguments.	become and remained well-informed. **Judiciousness:** the person provides great flexibility in considering alternatives and options to find more than one solutions, at the same time, someone possesses prudence in suspending when making or altering judgements.
	Conjecturing Alternatives: • to develop diverse solutions to problems and to predict possible outcomes.	
	Drawing conclusions: • to determine a position and an opinion on a specific issue by utilizing appropriate ways of reasoning	
Explanation	Stating Results: • to state the reasoning results accurately so that they can be analyzed, evaluated and monitored.	
	Justifying Procedures: • to present information about concepts, methods, standards and situations so that you or other people are able to examine the procedures of interpretation, analysis, evaluation and reasoning.	
	Presenting Arguments: • to provide the reasons for supporting arguments and respond to the objections to methods, concepts, arguments and standards accordingly.	
Self-regulation	Self-examination: • to reflect on the reasoning, examine their own opinions and arguments and rethink the factors which affect the fair judgement, such as knowledge limitations, stereotypes, basis, emotion, motivation, values and attitudes, etc.	
	Self-correction: • to discover problems which can be remedied and corrected in a reasonable way during the process of reflection.	

Appendix B. 21st century domains and competencies

(Condliffe *et al*., 2017, p.35)

Domains	Example competencies (knowledge and skills)
Cognitive Domain: Competencies related to thinking skills, such as reasoning, problem solving, and memory. This domain also includes content knowledge and creativity.	Academic content skills
	Critical thinking
	Technological literacy
	Active listening
	Problem solving
	Creativity
Intrapersonal Domain: Affective competencies used to "set and achieve one's goals" (Pellegrino and Hilton, 2012).	Self-regulation
	Metacognition
	Grit
	Flexibility
Interpersonal Domain: Competencies used to express, interpret, and react to information.	Communication
	Collaboration
	Conflict Resolution
	Leadership

Appendix C. Information for Understanding CCTST Score Report (34-point versions)

(Insight Assessment, 2020b, p.1)

C1. Qualitaive Interpreatation of Scale Scores

CT skills	Not manifested	Weak	Moderate	Strong	Superior
Overall	0-7	8-12	13-18	19-23	24-34
Analysis	0-2		3-4	5 or more	
Inference	0-5		6-11	12 or more	
Evaluation	0-3		4-7	8 or more	

C2. Qualitative Descriptions of Different Levels of CT Performance in Skills

Levels of CT performance	Qualitative descriptions
Not manifested	This result is consistent with possible insufficient test-taker effort, cognitive fatigue, or possible reading or language comprehension issues.
Weak	This result is predictive of difficulties with educational and employment related demands for reflective problem-solving and reflective decision-making.
Moderate	This result indicates the potential for skills-related challenges when engaged in reflective problem-solving and reflective decision-making associated with learning or employee development.
Strong	This result is consistent with the potential for academic success and career development.
Superior	This result indicates critical thinking skill that is superior to the vast majority of test takers. Skills at the superior level are consistent with a high potential for more advanced learning and leadership.

Appendix D. CCTDI numerical score range and qualitative interpretations
(Insight Assessment, 2021, p.24)

CCTDI score range and qualitative category	Recommended CCTDI Attribute Interpretation
Strong negative (10-19)	These scores are indicative of strong negativity or hostility toward the attribute being measured. Scores in this range indicate that the individual may systematically avoid situations requiring thinking and decision-making or work to disrupt or destroy decision processes where others seek to make strong judgments. Movement from this score range to positive score ranges is rarely observed, even after participation in an educational opportunity aimed at strengthening the specific attribute in question.
Negative (20-29)	These scores are indicative of poor valuation or aversion toward the attribute being measured. Scores in this range indicate negative habits of mind, and approaches to reasoning and problem-solving that are likely to limit the quality of thinking and decision-making and therefore diminish the fair-minded and willing engagement of problems in many situations. While movement from this level of score to one demonstrating strength on the attribute has been seen in some individuals after participation in an educational opportunity aimed at the specific attribute in question, it is more common to see an individual express at most an ambivalent valuation of the measure at the second assessment date.
Inconsistent/ Ambivalent (30-39)	These scores are indicative of ambivalent or inconsistent endorsement of the attitude or attribute being measured. Individual test takers are frequently seen to move from this score category to a higher range score as a result of completing an educational or training program aimed at developing this mindset attribute. More rarely, such efforts also result in some individuals regressing to a more adverse or hostile range at second assessment. This movement toward aversion or hostility is consistent with other observations of attitude formation. Having the habit of mind of engaging life and work problems with one's higher order thinking skills in order to determine what to believe and what to do (critical thinking), requires courage, persistence, honesty, and organization. It is reasonable that some individuals may reject the challenge or find this path too fearful or difficult. Many of these individuals fall back into non-reflective strategies to address life & work problems.
Positive	These scores indicate consistent endorsement and valuation of the

(40-49)	attitude or attribute being measured. Individuals scoring in this range typically demonstrate this mindset attribute with reliability.
Strong positive (50-60)	Scores in this range indicate that the attribute or attitude is a positive habit of mind and likely to factor into the individual's approach to all higher order thinking (reflective problem definition and problem-solving), particularly when the situation is of high consequence.

Appendix E. Questionniares

E1. The Effectiveness of CT Strategies Questionnaire on Improving Students' Critical Thinking (CT strategies questionnaire)

You are being invited to participate in a research study titled Critical Thinking Instruction Through Project-based Learning in Chinese EFL classes- A Case Study in HE to investigate the effectiveness of three critical thinking strategies (thinking maps, group discussion and peer review) on improving your critical thinking (CT) skills and dispositions. No one knows how effective these strategies in helping you improving your critical thinking better than you. It would be greatly appreciated if you would be involved in this process by completing the sheets attached, and returning them to me. The result of the questionnaire will be used only for research purpose. Please be as truthful as possible in completing the questionnaire and do not leave out any of items.

This questionnaire contains three sections. Please read instructions in each section and write your answers.

Thank you

学号_____ 性别： 男 □ 女 □

Part I

以下问题将要调查三种思辨教学策略（思维导图、小组讨论和同伴互评）对于提高你的思辨能力的有效性。请根据你本人了解的情况和个人感受，并按照问卷的要求选择在多大程度上（由 1-6 表示）同意或不同意下列条目。例如，“同伴互评提高了我的思辨能力”，如果你非常不同意，请选择 1。如果你非常同意，请选择 6。如果介于中间，请根据程度选择 2-5。

The following questions are about how agreeable you are with the effectiveness of three CT strategies, including thinking maps, group discussion and peer review, to improve CT skills. There are six scales to choose from 1-6. 1= Very disagreeable, 2= Disagreeable, 3=Slightly disagreeable, 4= Slightly agreeable, 5= Agreeable, 6=Very agreeable.

	Thinking maps helped me **思维导图能帮助我**	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
1	comment on the logical structure of the text with reference to the sequencing of the information. 通过信息的呈现顺序，帮助我评判自己作文的逻辑性。	1	2	3	4	5	6
2	conduct a comparative or constructive analysis of different viewpoints on the same topic. 对不同观点进行比较和分析。	1	2	3	4	5	6
3	evaluate whether the thesis statement is precise to the topic. 写作中心论点与主题相关。	1	2	3	4	5	6
	Group Discussion helped me **小组讨论能帮助我**						
4	collect evidence from different sources in support of an argument. 使我了解不同来源的证据，以便支持自己	1	2	3	4	5	6
	Peer review helped me **同伴互评能帮助我**						
5	evaluate whether the evidence or the reasons are relevant and precise to the claims. 使我反思我提出的理由或例子是否支持我的观点。	1	2	3	4	5	6
6	use cohesive devices (eg. comparison and contrast, cause and effect) to demonstrate relationships among sentences and paragraphs.帮助我更准确地使用衔接手段来表示对比、因果、递进等句子之间的关系	1	2	3	4	5	6

Part II

以下问题将要调查三种思辨教学策略（思维导图、小组讨论和同伴互评）对于提高你的思辨倾向的有效性。请根据你本人了解的情况和个人感受，并按照问卷的要求选择在多大程度上（由 1-6 表示）同意或不同意下列条目。例如，“同伴互评提高了我写作的自信心”，如果你非常不同意，请选择 1。如果你非常同意，请选择 6。如果介于中间，请根据程度选择 2-5。读完每个问题后，请给出你的即时的感觉，并在相应选项的数字上打勾。

The following questions are to investigate the effectiveness of CT strategies in increasing CT dispositions. There are six scales to choose from 1=Strongly disagree, 2=Disagree, 3=Slightly disagree, 4 =Slightly agree, 5=Agree, 6=strongly Agree.

	Thinking maps helped me **思维导图能**	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agreea
7	think deeply about the complexity of a phenomenon, a term or a concept. 使我思考什么样的证据能够真正支持自己的观点。	1	2	3	4	5	6
8	anticipate possible conclusions. 预见可能的结果	1	2	3	4	5	6
9	identify the central argument.理清哪一个是我最主要的观点	1	2	3	4	5	6
10	establish confidence in the reasoning process of my writing.对写作中推理出的结论自信	1	2	3	4	5	6
	Group Discussion **小组讨论能帮助我**						
11	inspired me to think in a different point of view. 使我从另一方面思考问题	1	2	3	4	5	6
12	reminded me to respect different opinions. 提醒我尊重不同观点	1	2	3	4	5	6
13	enabled us brainstorming more ideas.使我学到了新知识	1	2	3	4	5	6
14	rethink the structure of my writing based on my teammates' comments. 启发我重新思考并调整写作的结构。	1	2	3	4	5	6
15	be aware of the importance of using strong arguments to defend good argument. 使我意识到要使用有力的证据证明好的论点的重要性。	1	2	3	4	5	6

Part III

以下问题调查的是你对三种思辨教学策略在提高思辨能力有效性的程度感受。请根据你本人了解的情况和个人感受，并按照问卷的要求选择在多大程度上（由 1-6 表示）同意或不同意下列条目。例如，“同伴互评提高了我思辨能力”，如果你非常不同意，请选择 1。如果你非常同意，请选择 6。如果介于中间，请根据程度选择 2-5。读完每个问题后，请给出你的即时的感觉，并在相应选项的数字上打勾。

The following questions are to investigate the degree of your agreement with the effectiveness three CT strategies. There are six scales to choose from 1=Strongly disagree, 2=Disagree, 3=Slightly disagree, 4 =Slightly agree, 5=Agree, 6=strongly Agree.

		Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
16	Thinking maps enabled me to increase my CT skills. 思维导图使我提高了思辨能力。	1	2	3	4	5	6
17	Group discussion enabled me to increase my CT skills. 小组讨论使我提高了思辨能力。	1	2	3	4	5	6
18	Peer review enabled me to increase my CT skills. 同伴互评使我提高了思辨能力。	1	2	3	4	5	6

谢谢你的支持！

Date:_____

E2. The Effectiveness of PBL Questionnaire on Improving Students' Critical Thinking (PBL Questionnaire)

Dear student

You are being invited to participate in a research study titled Critical Thinking Instruction Through Project-based Learning in Chinese EFL classes- A Case Study in HE to investigate your attitudes and perceptions of PBL to improve your criticality and writing performance. This study is being done by Qiuling Bi from the University of Northampton. Your participation will help the researcher to gain a deep understanding of the teaching procedure of PBL on thinking skills and writing for college students. This is an anonymous survey, so there are no "right" or "wrong" answers. It would be greatly appreciated if you would be involved in this process by completing the sheets attached, and returning them to me. The result of the survey will be used only for research purpose. Please be as truthful as possible in completing the questionnaire and do not leave out any of questions.

Thank you

学号_____ 性别： 男 □ 女 □

Part I

以下条目是关于你对这门课程和其中教学环节的态度和观点，请根据你本人了解的情况和个人感受，并按照问卷的要求选择在多大程度上（由 1-6 表示）同意或不同意下列条目。例如，“我认为学习英语很难”，如果你非常不同意，请选择 1。如果你非常同意，请选择 6。如果介于中间，请根据程度选择 2-5。读完每条条目后，请给出你的即时的感觉，并在相应选项的数字上打勾。

The following items are about how agreeable you are with your own individual perceptions of and attitudes to the PBL class. There are six scales to choose from 1-6. 1= Strongly disagree, 2= Disagree, 3=Slightly disagree, 4= Slightly agree, 5= Agree, 6=Strongly agree.

	2. I think my teacher's feedback in writing could help me to revise my writing. 我认为老师给我写作上的反馈能帮助我修改自己的作文。	1	2	3	4	5
	3. I think it is ghlepful for me to gather inforamtion with my group members. 我认为小组成员一起搜集信息对于写作是有用的。	1	2	3	4	5
	4. I like to complete my writing project with the help and support from group members. 我喜欢从组员那儿得到支持和帮助来完成我的写作项目。	1	2	3	4	5
	5. I think PBL learning (writing) inspires me to concenrtate more on the real world around me, especially on the dialectic point of view. 我认为项目式学习（写作）启发我更关注周围的真实世界，特别是从辩证的角度了解或研究我感兴趣的话题（比如：爱情、亲情和个人成长等）。	1	2	3	4	5
	Your attitudes towards and perceptions of PBL					
	6. I think PBL improve my critical thinking. 我认为项目式学习提高了我的思辨能力。	1	2	3	4	5

Part II

Please share your suggestions about PBL teaching of this semester in English or Chinese.

_____

_____

_____

□ Please tick the box if you would like to accept interview.

Thank you for your support!

Date:_____

Appendix F. Interview Questions

1. Could you please describe how your teacher teaches English in high school?
（你能描述一下高中老师怎么教英语的吗？）

2. How about your writing learning?
（你是怎么学习写作的？）

3. Did your English teacher teach you something about critical thinking?
（高中老师有没有教过关于思辨的内容？）

4. How do you evaluate writing quality in high school?
（你在高中的时候作文的评分标准是什么？）

5. Does thinking maps improve your thinking skills? If yes, how did it help you?
（你认为思维导图提高了你的思辨能力吗？如果有的话是怎样帮助你提高的？）

6. Does peer review improve your thinking skills? If yes, how did it help you?
（你认为同伴互评提高了你的思辨能力了吗？如果有的话是怎样帮助你提高的？）

7. Does group discussion improve your thinking skills? If yes, how did it help you?
（你认为小组讨论提高了你的思辨能力了吗？如果有的话是怎样帮助你提高的？）

8. Which CT strategy do you prefer? Which one is the most effective? Why?
（哪种思辨方法你最喜欢？哪种是最有效的？为什么）

9. How do you evaluate your essay now?
（现在你作文的评分标准是什么？）

Appendix G. Interview Coding Lists

G1. An Initial Interview Coding List

※ The 1st interview question (students' English learning experiences in high school)
-Category 1: Evaluation of English class
Codes to be continued
-Category 2: Learning contents
Codes to be continued
-Category 3: Teachers' role
Codes to be continued
※ The 2nd interview question (How did students learn to write)
-Category 1: Ways of learning
Codes to be continued
-Category 2: How did they write
※ The 3rd interview question (Whether they learning CT in high school)
-Category 1: Never
Codes to be continued
-Category 2: Sometimes
Codes to be continued
-Category 3: Learnt a lot
Codes to be continued
※ The 4th and 9th interview questions (students' standards for evaluating essays pre-and post-PBL)
-Category 1: Standards before PBL
Codes to be continued
-Category 2: Standards after PBL
※ The 5th interview question (How thinking maps developed their thinking skills)
-Category 1: The writing development they made
Codes to be continued
-Category 2: The issues they were concerned about when drawing or revising thinking maps
Codes to be continued
※ The 6th interview question (How group discussion developed their thinking skills)
-Category 1: The writing development they made
Codes to be continued
-Category 2: The issues they were concerned about when taking part in group discussion
Codes to be continued
※ The 7th interview question (How peer review developed their thinking skills)
-Category 1: The writing development they made
Codes to be continued
-Category 2: The issues they were concerned about when peer reviewing essays
Codes to be continued

※ The 8th interview question (students' attitudes towards the effectiveness of each CT strategy to improve their thinking)

-Category 1: The most effective CT strategy.

Code 1: Thinking maps

Code 2: Group discussion

Code 3: Peer review

Code 4: They are the same

Code 5: reasons

-Category 2: Students' favourite CT strategy

Code 1: Thinking maps

Code 2: Group discussion

Code 3: Peer Review

Code 4: They are the same

Code 5: Reasons

G2. A Final Interview Data Coding List

※ The 1st interview question (students' English learning experiences in high school)
- -Category 1: Learning contents
  - Code 1: Handwriting
  - Code 2: Linguistic knowledge
- -Category 2: A writing format in the exam
  - Code 1: The given writing outline
- -Category 3: Teachers' role
  - Code 1: A leading role

※ The 2nd interview question (How did students learn writing)
- -Category 1: Ways of writing learning
  - Code 1: Memorsing the model essays
  - Code 2: Teacher's writing lectures
- -Category 2: How did they write
  - Code 1: Following the given outline

※ The 3rd interview question (Whether they learning CT in high school)
- -Category 1: Never
- -Category 2: Sometimes
- -Category 3: Learnt a lot

※ The 4th and 9th interview questions (students' standards for evaluating essays pre-and post-PBL)
- -Category 1: Standards before PBL
  - Code 1: language
    - Code 1.1 Sentence patterns
    - Code 1.2 Advanced vocabulary
    - Code 1.3 Grammar
  - Code 2: Handwriting
- -Category 2: Why students focused on language instead of thinking in high school
  - Code 1: They were given a writing direction with three points
  - Code 2: Teachers' teaching focus
    - Code 2.1: Linguistic competence
    - Code 2.2: Handwriting
- -Category 3: Standards after PBL
  - Code 1: Language
    - Code 1.1 Advanced vocabulary
    - Code 1.2 Sentence patterns
    - Code 1.3 Grammar
  - Code 1: Critical thinking
    - Code 1.1 Organisation
    - Code 1.2 A clear thesis statement
    - Code 1.3 A good structure

Code 1.4 Logic

※ The 5th interview question (How thinking maps developed their thinking skills)

-Category 1: Thinking skills' development

Code 1: Development of the Interpretation skill

Code 1.1 The writing development: a clear mind

Code 2: Development of the Analysis skill

Code 2.1 The writing development: more oraganised writing

Code 2.1.1 Increased awareness of writing sequence

Code 2.1.2 Enhanced capacity of determining writing focus

Code 2.1.3 Improved consciousness of coherence

Code 2.1.3.1 Focus on the thesis statement

Code 2.1.3.2 Not miss key points

Code 3: Development of the Inference skill

Code 3.1 The writing development: proper examples to support ideas

Code 3.2 The writing development: following a reasoning sequence

-Category 2: A benefit besides writing: higher writing speed

-Category 3: The issues they were concerned about

Code 1: Contents of the thinking map

Code 2: The difficulty to visualise the ideas

Code 3: The confusion in applying the maps' contents to writing practice

※ The 6th interview question (How group discussion developed their thinking skills)

-Category 1: Thinking skills' development

Code 1: Development of the Evaluation skill

Code 1.1 Preparing for writing: detecting errors

Code 1.2 Preparing for writing: brainstorm all possible ideas

Code 2: Development of the Interpretation skill

Code 2.1 Preparing for writing: brainstorm all possible ideas

Code 2.2 The writing development: more comprehensive writing

-Category 2: The issues they were concerned about

Code 1: Less active participation

Code 2: Less effective discussion time

Code 3: Dissatisfaction about gains from discussion

※ The 7th interview question (How peer review developed their thinking skills)

-Category 1: Thinking skills' development in writing

Code 1: Development of the Analysis and Evaluation skills

Code 1.1: Being reviewers: Error detection

Code 1.1.1 Language correction

Code 1.1.1.1 Grammar points

Code 1.1.1.2 Vocabulary

Code 1.1.1.3 Sentence patterns

Code 1.1.2 Critical thinking correction

Code 1.1.2.1Coherence

Code 1.1.2.2 Linking devices
Code 1.1.2.3 Logical issues
Code 2: Development of the Self-regulation skill
Code 2.1 Being student writers: Self-examination and self-correction
-Category 2: Being reviewers: More opportunities to learn from others
Code 1: Linguistic development
Code 2: Avoiding making the same mistakes as the writers
-Category 3: The factors negatively influenced peer review
Code 1: The limited capacity to recognise mistakes
Code 2: Inactive participation
Code 3: Anxiety about hurting other people's self-esteem

※ The 8th interview question (students' attitudes towards the effectiveness of each CT strategy to improve their thinking)

-Category 1: The most effective CT strategy.
Code 1: Thinking maps
Code 2: Group discussion
Code 3: Peer review
-Category 2: Students' favourite CT strategy
Code 1: Thinking maps
Code 2: Group discussion
Code 3: Peer Review

Appendix H. Example of Transcription of Interview

Chinese version

毕：你高中的的那个英语的学习经历是怎样的?

S8：经历？

毕：就是你上课老师怎么讲的？

S8：嗯。一般是上课大概有一半到 2/3 的时间是讲前一天晚上的作业。

毕：作业？

S8：就是，老师会每天给我们布置作业，他有自己那种规划，就是按照高考的题型，哪段时间应该做哪种题型，然后剩下一点时间讲课文。

毕：老师讲的重点东西是什么？哪一方面呢？

S8：嗯，语法吧。

毕：还有啥别的吗？

S8：嗯，文章的框架。就作者思路啊什么的。基本上不讲单词

毕：那写作方面老师怎么讲呢？

S8：好好想一想啊。大概每一次月考的时候作文它会重点的讲，就是分析题目呀，然后哦哦，他有一套自己作文的，一个固定的怎么说也不能叫模板，就是写几段啊，每段写多少字啊什么的。

毕：每段多少字都有啊。

S8：不是，就是整篇文章他要求是 100 几，我忘了。他让我们写 105 到 110 嘛，不能多也不能少啊，然后写三段，然后开头、中间和结尾怎么怎么样。哦，还有就是在临期末考试前一段时间会让我们背谚语什么的。

毕：背谚语，范文有吗？

S8：哦！也有。就到了高三的时候，给我们就是一整套就有好多篇，然后让我们每天早读的时候背，然后检查那样。

毕：所以就说老师帮助你们就是提高写作能力的方式，一个是背范文、背谚语，然后再就是老师会讲你说的那种什么三段式啊什么之类的是吧？那你们有没有整个高中阶段就不光高三，还有高一高二，有没有上过单独的写作课呢？

S8：我没上过。就老师没有给我们开过。

毕：就简单讲过写作是吧。

S8：嗯，有。单独是有的。他是每一个题型都会有个 PPT，就是试卷中的题型。嗯，然后在他认为合适的时候就抽一节课来给我们讲这个。

毕：啊，那你们讲了会写吗？

S8：不写就是光讲，

毕：那这个写作，在你印象当中这三年里边老师大概单独讲过几次？比如说一个月一次？

S8：不不不，没有那么频繁。一个学期有一次就差不多，就很好了。我跟了他三年，一直他是他教我，然后也就讲过两三次吧。

毕：那你能就是简单的描述一下，你（对英语学习的）的感觉嘛？

S8:我觉得我的英语成绩其实不算太好，只是分数在老师的帮助下虚高，实际英语水平不太好。实在要说的话，就是英语大部分不是死记硬背出来的，很多语法都是在阅读课文和阅读

理解的时候自然而然领悟记住的，就像小时候的语文学习一样，读很多书自然而然就学会阅读写作了）但是我们练字练得很多。

毕：好像好几个同学都提到过这个问题了，大家都好重视。

S8：就是好像确实字好看了，分数上个档次。

毕：你的作文分数是多少？你的总分呢？

S8：写作不清楚。总分 135。

毕：那你觉得咱们现在这种上课模式和以前有什么不一样的地方吗？

S8： 高中的时候老师也会引导我们去思考，每节课留出一半甚至更多时间来让我们去讲自己找到的知识点。但是这种思考是被动的，也仅限于知识层面。大学里老师讲的很少，主要是小组讨论，我们站的位置也更高了，是自己主动去思考文章架构等等。）

毕： 还有吗？

S8：嗯。有吧。感觉，嗯...就以前的课好像是练的比较少，想的比较多，后咱们现在的课大概是感觉就练的比较多，然后,啊。

毕：那老师呢？

S8：嗯。（停顿）

毕：我的意思是说老师在课堂当中你觉得她的作用没有发生变化呢？

S8：嗯。（停顿）感觉上大学以后就对四级考试的，对考试试卷几乎就不是很了解。啊，就是老师几乎就不会去讲这种东西。然后在高中的时候，老师是不怎么按课本的。就是完全，就几乎是按照考试的那个来的。

毕：对，是因为四六级考试并没有你们高考那么重要，因为高考那个事情是非常重要的，我觉得应该不光英语（课），别的课应该都是围绕着高考吧，是吧？

S8：对，几乎所有的课。嗯，就是高中是可以几乎不看课本的，有一些科目。但是大学就看课本就可以了。

毕：那你觉得现在是啥样的模式呢？

S8：嗯。就很很很考验自己的自律吧，我感觉。自律啊，就是呃...没有说那种就是强制性的，你不做就受什么惩罚，不会有那种东西。就是如果你要想提高的话，你就在课下自己好好学，但是你要是想混水摸鱼的话其实也挺容易。然后就全靠自己学的感觉啊！

毕：嗯，所以你就说老师原来可能是起一个，比如说在上课时候是主导的，就是大家都按照我这个来就行，那所以现在就变成了一种。所以可能现在就变成了一种，你觉得现在老师的作用是啥呢？嗯。

S8：不知道（笑）

毕：很坦诚。（笑）有的同学想不出来就一直在那儿憋着。我说你可以想想我说的有没有道理哈，老师是不是在旁边就是起一个，就是，辅助的一个作用？就是老师不是监督你去做，就主要是你们去做。然后老师可能在旁边就起一个

S8：指路。

毕：对，引导，嗯，是吗？那就是一个引导的一个作用哈。

毕：那咱们这个整个上课过程，这个你觉得这个思维导图对你写作有用嘛？

S8：嗯。有用吧。嗯，要是没有思维导图的话可能就就像是很多句子摞在一起，然后如果哪一天你摞不起来了，你想不出来句子这个完了，你就写不完了，或者你不知道该写什么了。然后有了这个架，至少你知道就那每一个，那种小主题上稍微写几句话，你这篇文章就起来了，就感觉还挺有条理。

毕：嗯，就是一个是有条理是吗？这就是可能有点逻辑，像你说的，如果我原来只是句子、句子，摞句子。写到后面我就不知道写啥了，但是我有了一个框架，我知道要围绕这个主题写这个句子是吧？

S8：嗯，

毕：对，是，嗯。那小组讨论呢？

S8：有，就是嗯，就课前会稍微想一想，就是大概确定一个什么题目，然后在课上会说一说，然后大家会有话题，哪个比较好就选哪个，但是不会说的那种，比如说写一篇文章啊，或者是嗯不会特别去隆重的准备啊

毕：你们组一直是这样做的吗？

S8：嗯。

毕：我看到有的组好像就不是很积极。

S8：我们组有一个同学不是很积极哦，然后剩下四个基本上就会说自己的观点，然后大家就互相商量商量呗。

毕：为什么她不参与呢？

S8：她说自己英语特别不好，可能是因为有时候听不太懂所以排斥？我也不太清楚

毕：嗯，大家都在讨论，他不说话自在吗？

S8：感觉无所谓吧就好像那种，反正我就是不会啊，然后你们说什么我也听不懂。

毕：你们有没有想过怎么让她融入到你们的谈话中呢？比如特意问问她的看法？

S8：其他组员不清楚，我个人没有过这种想法。同学们的态度应该基本都是你自己的事自己看着办，多说了反而显得自己很奇怪，除非是关系很好的朋友才会劝说帮助。

毕：那你觉得你们小组讨论对这个写作主题的确定有帮助吗？

S8：当然有啊。嗯，就你一个人想的话，将来这个题目可能...嗯，就是，我们就是选一个最能说的题目吧。啊，至少是自己想的，可能会有一些比较偏的地方吗。

毕：那你们组里决定写作题目的标准是啥呢？

S8：哦，就是大家觉得这个题目既符合这个单元的主题，嗯，然后又能够引申出来很多。就你只要一想这个题目，你就可以直接有一个很完整的，就能想出来一个框架那种啊。就有一些有些题目，你一想你是想不出来那种非常清晰的框架的，就不太好写。

毕：嗯，好，那你们题目确定之后，比如说大家认为我这个写的比较好，然后我想的题目为标准画一个思维导图，大家帮我改，还是说我们一块儿来做这个思维导图呢？

S8：一块儿。

毕：嗯，那你觉得这样的方式对你的这个思维有什么帮助吗？就整个这个过程里边。

S8：嗯。有吧。就是思路可开阔。

毕：你能具体说说吗？

S8：思路更开阔就是它能够，嗯，就是能够引申到很多方面去。

毕：刚刚你说刚刚一开始在写的时候不知道问题的缺陷在什么地方，她（组员）提出更好的意见，让你认识到了你这个里边的问题，那你的问题是什么？

S8：嗯，就是...写出来就很空。

毕：你说的空的意思是？嗯，就是这个话我在这个话题里边儿说合适，在别的话题里边儿说也合适，还是说，我只是写了这一句，然后我后面没有东西了。这个空是指的是哪一种呢？嗯。就具体的例子，你还能想起来吗？

S8：嗯。就是，可是没有没有逻辑吗？就是我说完了写完这句话以后，我可能就不知道接下来该说什么了，但是如果从就是他那个题目的话，我写完了这个，我就知道接下来该写什么那种感觉。

毕：哦，就是你思路开阔的意思就是逻辑性，是吗？

S8：逻辑性啊，大概是我说这话也很没有逻辑（笑）

毕：那如果说你一个人主笔，然后大家都帮你改，也帮你弄这个思维导图，然后帮你修改作文和那种我自己写一篇，你觉得哪一种比较好一些呢？

S8：嗯，大家改吧。

毕：大家改，嗯。你觉得大家改的好处有啥呢？

Interviewer: What benefit does it have?

S8：就嗯，有一些地方就逻辑不通的地方，或者是嗯，比如说就是有些地方就别人看我（文章）感觉就像缺一块，可能还需要一些，比如说关联词或者是可以加一段什么东西，但是自己可能是看不出来的。嗯，就可能语意更通顺或者是文章就更完整。

毕：语言方面有帮助吗？

S8：嗯，有吧。就是上次我看到组里同学写的，她的逗号用的非常奇怪。感觉逗号都是中文的用法。就是跟英文的逗号是完全不一样的用法。这种东西她自己可能也看不出来。

believe she did not realise it.

毕：因为他思维就是汉语的思维，你看他写的是英语，其实他思维就是汉语的，所以它那个逗号像你说的，就是汉语的那种。

S8：可能是就一句话里面三四个逗号。

毕：这种项目式学习这个过程吗？

S8：嗯，感觉更自由吧。

毕：更自由，你觉得这个自由表现在什么方面呢？

S8：哦！就没有那种就被被老师牵着走，你说往这儿走，我必须跟着你往这儿走，就是你可以自己去思考一些东西。

毕：那你思考的东西是什么呢？

S8：嗯...（停顿）

毕：按照这种模式学习，这一学期的学完，你的感受是什么？就像你前面说的，高中学习，老师有自己的一套，她很清楚，就是计划性很强，学生按照坐做就行。像你说的很自由，表现在什么方面呢？

S8：就是学习的方式吗？因为以前做完老师布置的作业就没有什么时间了。这学期我就可以利用自己的时间，背背英语课文啊，或听听 BBC 什么的。就能够通过自己想要的方式学习英语。

毕：就是学习的方式自由了。再就是我们这些环节。比如说思维导图啦、小组讨论啦，同伴互评啦，这些环节中你比较喜欢哪个环节呢？

S8：讨论吧。

毕：为什么呢？

S8：那个时候，叫什么？思维的碰撞。就是你能听到不同人的看法。面对某个问题，她和你不一样的地方在哪里，原来还可以这么想。

毕：通过集体讨论，有没有感觉自己的想法更开放了，或者说表达自己的观点更自信的表达了。

S8：在自由表达这方面，就大家挺认同我的一些看法的。

毕：你觉得你的想法跟别人有什么不一样的地方，让大家比较认同你的想法？

S8：可能是我的语言表达能力强一点吧，可以更清晰准确的说出我的想法，有些同学想法可能很好，但是没有办法让别人理解。

毕：这几个环节里哪个比较有挑战性？

S8：没有吧。

毕：就是都能完成？

S8：嗯。

毕：那个评分标准你们用过吗？

S8：没有。

毕：嗯，为什么没用呢？

S8：不知道为什么每次都忘了它的存在。

毕：你自己对英语学习的期望值是怎么样的呢？或者说是英语学习的目标是什么？

A:其实期望值蛮高的，希望自己可以达到用英语流畅的表达自己，不管说还是写，也可以自如的阅读外文文学、期刊、论坛等等。）

S8：其实考研的时候，大学要用英语面试。

毕：对，是。

S8：好难。愁死了，

毕;那还有好几年，嗯,我觉得那个是可以慢慢积累的。

S8：是的。

毕:经过一学期的学习,如果现在让你评价一篇文章的好坏(可能是任何你看到的英语文章),应该从哪几个方面评价？这几个方面里按照重要性排序是怎么样的？

S8：1.文章表达的核心内容，抒情议论类的文章看它表达的价值观等等，科普类的一般是介绍人们不太了解的知识，都挺有价值的。2.文章结构安排，我认为一篇文章结构清晰详略得当才能正确表达出作者想要表达的东西，否则写了也会让人抓不住重点。3.语言表达能力，合适的修辞词汇才能表达出文章正确的力度，才能传达作者的情感或者实现文章的精确性。

毕：如果大学现阶段的作文重点是在语法、语言上，这和中学时候的英文写作有什么区别？或者说理论上，大学阶段的英语学习（在写作方面）应该比中学阶段学的更深一些，这个深度表现在什么方面？

S8：一方面是词汇量的增加，一方面是更深的逻辑性增强，文章更有层次感。

English version

Interviewer: Could you please tell me about your English learning experience in high school?

S8: Experience?

Interviewer: How did your teacher conduct her English class?

S8: Well, she spent half or two-thirds of the class time giving a lecture on the homework last night.

Interviewer: Homework?

S8: Yes. They are different types of questions required in the English examination in Gaokao. So the teacher gives us similar questions as homework to prepare us for Gaokao. The teacher will explain the article in the textbook for the rest of the class.

Interviewer: What was her focus?

S8:Grammar.

Interviewer: Anything else?

S8: The outline and storyline line of the article. She seldom discussed the vocabulary.

Interviewer: What about writing?

S8: Silence

Interviewer: Any impression?

S8: Let me think about it. Almost every month, nearer the monthly examination, she gave us a lecture on essay writing, such as how to analyse the topic. She had her template. For example, how many paragraphs we should write and how many words we should write for each paragraph.

Interviewer: Even the number of words in each paragraph?

S8: Not exactly. We were required to write 120 words in Gaokao. She asked us to write anything between 105 words to 110 words. Then we had to write three paragraphs and what we should write at the beginning, middle, and end of the essay. She also asked us to memorise some proverbs.

Interviewer: Proverbs. How about model essays?

S8: Oh, yes! She gave a whole set of sample essays, including several pieces when we were in Grade 12. She also asked us to memories them and when would check on us..

Interviewer: So the methods for English writing introduced by your teacher are to memorise the proverbs and to recite model essays, and the three-paragraph structure. Did you have any writing course?

S8: No.

Interviewer: But she had such lectures.

S8: Yes. Actually, she had PPTs for each different type of questions, including writing.

Interviewer: Ok. Did you practice writing after the lecture?

S8: No. She only gave us the lecture .

Interviewer: How many times did she give the writing lecture? Every month?

S8: Not so often. I remembered two or three times for all three years in high school.

Interviewer: Could you please describe how do you feel about English learning in high school?

S8: I don't think my English level is that good. I just got a high mark in the examination with the help of my teacher. If you really wanted me to share something with other students, we can never learn English well by memorising. The more you read widely, the more you get from reading comprehension. This is like the way we learn Chinese, and we learn how to write in Chinese by reading a lots of books. We also practiced handwriting a lot!

Interviewer: Several students mentioned the importance of handwriting!

S8: Our score becomes higher if we have good handwriting

Interviewer: How about your score in writing and final score?

S8: I have no idea about my writing score. My final score is 135。

Interviewer: What is the difference in the class mode between the high school and the university?

S8: When I was at the high school, the teacher also encouraged me to think, and she would leave half of the class time or even more time for us to find the knowledge points. But this was a kind of passive thinking, and it was limited to the knowledge level. Right now, in the university, teachers lecture less, and we have more group discussion. We stood in a high position and learn more actively about how to organise an article, for example.

Interviewer: Anything else?

S8: Emmm. I think we had less prarice in writing in high school and we had more writing practice right now.

Interviewer: What about the teacher?

S8: (pause)

Interviewer: Is there any difference in the teacher's role in class between high school and the university?

S8: Emm

Interviewer: As you said before, the teacher had her plan and students follow that plan to learn.

S8: (Pause) At the university, I think I am not familiar with the CET-4 examination because the teacher rarely mentioned it. While in high school, my teacher did not follow the contents in the textbook but her teaching is always oriented towars Gaokao.

Interviewer: Well, that's because CET-4 is not so crucial as Gaokao. It is so critical that all the courses, including English, took Gaokao as the central task.

S8: Yes, in high school almost all the courses, we did not even need to read the textbooks While in the university we just learnt the textbooks.

Interviewer: What about now?

S8: Well, self-discipline is an important skills for university study test of, I think. After all, you won't be punished if you do not do the assignment. If you want to improve yourself, you have to make great efforts after class. However, if you want to

muddle through life, you do not have to study hard. It depends on yourself. This is unlike the teachers in high school. They are like nannies who take care of you all day long.

Interviewer: So the teacher played a leading role in class. What do you think the role the teacher played in our class?

S8: I have no idea (Smile)

Interviewer: You are straightforward. (Smile) You can consider whether it is sensible. Does the teacher play an assistant role? I mean, the teacher does not supervise students. Instead, students decide what to do to achieve the learning outcome. So the teachers may lead a ...

S8: Directing the way.

Interviewer: Ok. A guiding role.

S8: But it's hard to improve English level following the same way. Actually, I don't know how to improve. Even if I got a high score in the examination, it was still hard to communicate. It's hard to tell. Sometimes, even though you can memorise the whole book, you were unable to improve your grade; sometimes your grade is improved, but it doesn't mean that your oral communication skills are good.

Interviewer: Like some contexts in listening class?

S8: Yes.

Interviewer: Do you think it is helpful to use thinking maps in writing?

S8: Well, it helped. If there is no thinking map, it seems that several sentences are bounded together unless you have no idea what to write next. If I have this map, it appears that I have a framework to write, and at least I can write few sentences to each branch and sub-branch. I feel the writing is organised in structure.

Interviewer: Oh, well organised. This is about logic. As you said, I originally wrote sentence by sentence, and at last, I am running out of word. If I have a framework, I can write more sentences around the thesis.

S8: Yes.

Interviewer: Ok. What about group discussion?

S8: Each of us thought about a topic before class and shared it witin the group. We then chose the best topic, but we didn't share information such as how to prepare an essay.

Interviewer: Did your group always follow the same procedure?

S8: Yes.

Interviewer: Wow. What I observed was some groups just decided on the writer before they discussed it.

S8: There was one teammate who did not participate actively. The rest four of us discussed, shared and decided on the topic.

Interviewer: Why did she not participate?

S8：She said her English was not good. Maybe she didn't understand when discussing it. I am not clear.

Interviewer: Was she embarassing?

S8: I think she doesn't care. I just can't do it when I don't understand what you are talking about.

Interviewer: Have you ever considered how to help her get into your discussion. For example, did you ask her opinion on purpose?

S8: Personally, I never thought about it. I think classmates' attitudes were it is ok that you mind your own business. But it's weird if you said too much unless they were close friends of yours.

Interviewer: Do you think group discussion helped you to determine the title of writing.

S8: Of course, it helped. If you thought about a title by yourself, it might be too narrow. But at least we can decide on a title together.

Interviewer: What is your group's standard to select a topic?

S8: Well, the standard is as long as the title is consistent with the topic of the unit and can be extended to many aspects to write. As long as you think about the title, you can draw a complete framework in your mind. This does not happen in every title.

Interviewer: After deciding the title, how do you draw the thinking map. Did your group draw the map by the writer or by the whole group?

S8: Together.

Interviewer: Do you think working as a group was helpful in thinking?

S8: Yes.

Interviewer: For example, it will be more comprehensive if I listen to other people's opinions. Or they provided a new angle that I never imagined, and my thought seems partial.

S8: Well, my mind is open.

Interviewer: Could you please say more?

S8: It means that my mind can be extended to include more aspects.

Interviewer: Just like what you said, you don't know your drawbacks in writing. You realised the problems when your group members made better suggestions. What were your problems in writing?

S8: What I wrote was empty.

Interviewer: What do you mean by saying "empty"?

S8: Well, maybe it was a lack of logic? I don't know what to say after writing a sentence. But if I use the title from other group members, I know what to write.

Interviewer: So your open-mindedness means the logic, right?

S8: Maybe I am not logical by saying such a statement.

Interviewer: Which one do you prefer if you wrote and revised all by yourselves or you wrote at first, and then the other people revised the writing.

S8: I prefer the second one.

Interviewer: What benefit does it have?

S8: For example, my teammates found one place that seemed incomplete, and I needed to add a linking word or one paragraph to explain further. I may not find such a place by myself. My essay will be more consistent or more complete following their suggestions

Interviewer: Complete structure, consistent content. Anything else? What you said was about a thinking line. What about the language?

S8: Well, it has. Actually, some students wrote in Chinese at first and translated it into English by software. But some translations are weird to be translated into English correctly.

Interviewer: I found this situation, especially for students whose writing is challenging to read in English. It is more reader-friendly if you translate it into Chinese.

S8: I found the comma usage is quite strange in one of my teammates' writing. I think her way to use the comma was the Chinese way which is quite different from English. I believe she did not realise it.

Interviewer: That's because her thinking in writing was Chinese, not English thinking. She wrote an English essay in Chinese thinking.

S8: She used three or four commas in one sentence.

Interviewer: What do you think of this writing project? Do you like this process?

S8: I think it gave us more freedom.

Interviewer: What aspect do you think is freer?

S8: Oh, we don't have a feeling of being passively followed by the teacher's instruction. You can think by yourself rather than being led to the direction that teachers gave you.

Interviewer: What is your consideration?

S8: Well... (pause)

Interviewer: How do you feel following this way to learn. As you discussed before, it was easy for students to follow teachers' instructions in high school because they are professional and well-planned. Now you are free. What freedom do you have?

S8: The way we learn? This semester I had more time to myself, so I can use my method to learn English, such as memorising the passage in the textbook or listening to the BBC radio. While I rarely have my own spare time because teachers gave too many assignments in the past.

Interviewer: Which one do you prefer, thinking maps, group discussion and peer review?

S8: Group discussion.

Interviewer: Why?

S8: What is it called? Brainstorming. I can hear different opinions from different students. They are so different from me towards the same issue. There is another way to think about it.

Interviewer: Which one do you think is challenging?

S8: No.

Interviewer: So you can do all these activities.

S8: YES.

Interviewer: Do you use the rubric to evaluate your writing?

S8:No.

Interviewer: Why?

S8: I don't know why I always forgot it.

Interviewer: Your revision focused on the language, like grammar and vocabulary, or on the logic, or using proper examples if they have these problems.

S8: Basically the language. We also commented on the problem of the structure in writing. But not every essay had this problem.

Interviewer: What is your expectation of English learning? Or what is your learning target?

S8:Very high. I hope I can express my opinions fluently in written English. And I can read English novels, journals and forums.

S8: We have to attend an interview as part of the postgraduate entrance examination.

Interviewer: Yes.

S8: It's so hard.

Interviewer: It takes you years. I believe you can improve your abilities gradually.

S8: I agree with you.

Interviewer: How do you evaluate the quality of an article after this semester learning?

S8: I think the most important one is the core content and value that the writer aimed to express. For example, I find it valuable to learn about the unknown knowledge in general science articles. The second one is the structure. I believe writers express their opinions clearly with a clear organisation and emphasis. The last one was the language. The article is reliable and displays the writer's true feelings with accurate expression.

Interviewer: Which aspect do you think you should improve if the writing in the university should be more profound than in high school?

S8：One aspect is to increase vocabulary. The other aspect is to enhance the logic of the article.

Interviewer: What development do you make after one semester's learning?

S8: I think my writing is more logical. I will prepare a structure then add more contents according to the structure. .

Appendix I. A rubric for assessing CT skills in writing

Core CT-skills	Items	Acceptable	Weak
Analysis	1. Elaborate on the thesis statements	State a specific writing plan to discuss or prove	No writing plan or only a vague statement to discuss or prove
	2. Provide main claims	Provide relevant key points to support the thesis statement	No relevant key points or the provided key points fail to support the thesis statement
	3. Provide sub-claims	Provide relevant statements to support main claims or upper-layer sub-claims	The statements are irrelevant or fail to support main claims or upper-layer sub-claims
	4. Compare and contrast	Identify and analyse differences or similarities of things compared or contrasted	Fail to display differences or similarities of things compared or contrasted
Inference	5. Collect evidence from different sources in support of claims (personal experience, statistics or citing sayings)	The evidence from different sources supports claims or has a close relationship with claims	The evidence from different sources fails to support claims, or thas no relationship with claims
	6. Use cohesive devices to demonstrate and construct logical and coherent relationships	Provide convincing devices to display the logical relationships between sentences or paragraphs	The devices fail to display the logical relationships between sentences or paragraphs
	7. Draw logical conclusions	Draw a logical conclusion based on the judgment of the previous information	No conclusion is provided, or draw an overall summative conclusion based on a summary of the previous information
Evaluation	8. Assess whether claims or evidence are workable, valuable or convincing	Accurately evaluate the feasibility or reliability of the statements (claims or evidence)	Lacking evaluation or the evaluation has no relation to statements (claims or evidence)
	9. Express implications	Accurately discuss possible effects or results based on the	Lack of evaluation or failure to discuss possible effects or

		statements	results based on the statements
	10. Determine what needs doing in a given situation	what is suggested to do can solve problems or make progress in the future	No statements or what is suggested to do has no relationship with or is too vague to solve problems or make progress in the future

Appendix J. Consent Forms

Please read each statement below and then confirm that you agree or disagree by placing your initials in the appropriate box.

	Yes	**No**
I have read and understood the information provided to me in the participant information sheet and the researcher has explained the research project to me clearly.		
I have had an opportunity to ask questions about this research.		
I agree to the interview being audio recorded.		
I understand that I can decline to answer any questions.		
I understand that my participation is voluntary and If I decide to withdraw, I can do that before the researcher make the data anonymous.		
I understand that I will NOT be affected by my personal comments I made about this research.		
I agree to anonymised quotations being used in any academic presentations or publications of this work.		
I understand that all the data will be kept confidential.		
I agree to my data (anonymised) being used in any subsequent work (i.e. PhD, related publications, conference papers and evidence to conduct a teaching reform) that builds on this current project.		
I agree to take part in the above study.		

Signature and date of person giving consent (the participant).

Signature and date of person obtaining consent (the researcher).

Appendix K. Ethics Lists

Appendix K1. Participant Information Sheet for Interview

Participant Information Sheet (interview)

Research Title

Critical Thinking Instruction Through Project-based Learning in Chinese EFL classes- A Case Study in HE

Why have I been invited?

You are being invited to take part in this research study on 12/21/2019. Before you decide whether you wish to participate, it is important for you to understand why the study is being conducted and what it will involve. Please take some time to read the information provided and discuss it with others if you wish. Please ask if there is anything that is not clear, or if you would like more information.

What is the purpose of the study?

The aim of this study is to gain a deep understanding of your attitudes towards and perceptions of the impact of PBL on developing criticality and writing.

Why have I been chosen?

You have been asked to take part in the study because you will be undertaken the program of College English in which some innovative approaches are included. Your opinion will enable this approaches to be evaluated.

Do I have to take part?

Taking part is entirely voluntary. If you decide to take part, you will be asked to sign a consent form to confirm that you understand the project and are willing to participate. If you decide to withdraw, you can do that before I make the data anonymous.

What will my participation involve?

Once you have agreed to take part, you will be individually interviewed for about half an hour by the researcher at the end of this semester to share your attitudes and perceptions of the course.

We will send you the text of any summary of your interview before using it in any project publications or reports.

What are the possible benefits of taking part?

The information obtained from this study will be used to complete this research project, and then will be publicised through journal articles. Reflection on participation will also enable a more in-depth understanding in application of project-based learning to teach English as a foreign language.

What are the possible risks or disadvantages of taking part?

You may feel worried about your negative comments may affect your grades on your intellectual performance. However, as a part of research data to explore effectiveness of project-based learning, you will not be assessed by your attitudes or perceptions in this research.

What if something goes wrong?

If you have any concerns about any aspect of the way you have been approached or treated during the course of this study, then please contact the researcher at the University of Northampton, Qiuling Bi (Qiuling. Bi@northampton.ac.uk).

Will my information be kept confidential?

All the information collected for this study will be anonymised and stored securely on Onedrive of the researcher's university, which is a secure university online repository with log ins which are controlled by the team of this research project. This repository allows us to retrieve the data at any time, but prevents unauthorised access.

What will happen to the results of the study?

The results from this study will be used in the following ways: related publications, and conference papers. If you would like to receive a summary of the findings, then please indicate this on the consent form or contact the lead researcher.

Who has reviewed the study?

This study has been reviewed and approved by the University of Northampton Research Ethics Committee on [date].

Contact for further information

If you have any questions about this study or your possible involvement then please contact me using the contact details below.

Qiuling Bi, the researcher, (Qiuling. Bi@northampton.ac.uk)

Dr. Qian Zhang, PhD Student Supervisor in the Faculty of Education and Humanities, (Qian.Zhang@northampton.ac.uk)

Thank you for considering taking part in this study.

Appendix K2. Participant Information Sheet for Questionnaire

Participant Information Sheet (questionnaire)

Research Title

Critical Thinking Instruction Through Project-based Learning in Chinese EFL classes- A Case Study in HE

Why have I been invited?

You are being invited to start to take part in this research study on 09/09/2019. Before you decide whether you wish to participate, it is important for you to understand why the study is being conducted and what it will involve. Please take some time to read the information provided and discuss it with others if you wish. Please ask if there is anything that is not clear, or if you would like more information.

What is the purpose of the study?

The purpose of this study is to investigate your attitudes and perceptions of effectiveness of project-based learning class to improve critical thinking and writing performance.

Why have I been chosen?

You have been asked to take part in the study because you will be undertaken the program of College English in which some innovative approaches are included. Your opinion will enable this approaches to be evaluated.

Do I have to take part?

Taking part is entirely voluntary. If you decide to take part, you will n to sign a consent form to confirm that you understand the project and are willing to participate. If you decide to withdraw, you can do that at any time without given any reason.

What will my participation involve?

Once you have agreed to take part, you will be asked to complete totally six questionnaires in class. One is a PBL class evaluation, which you will be complete after finishing this course. The other five are regarding your performance evaluation

in group discussions, which you will complete separately in weeks three, six, nine, twelve and fifteen immediately after class.

What are the possible benefits of taking part?

The information obtained from this study will be used to complete this research project, and then will be publicised through journal articles. Reflection on participation will also enable a more in-depth understanding in application of project-based learning to teach English as a foreign language.

What are the possible risks or disadvantages of taking part?

We do not believe that there are any possible risks. However, we will monitor data as it is collected and discuss implications of any risk which emerge with the participants involved.

What if something goes wrong?

If you have any concerns about any aspect of the way you have been approached or treated during the course of this study, then please contact the researcher at the University of Northampton, Qiuling Bi (Qiuling. Bi@northampton.ac.uk).

Will my information be kept confidential?

All the information collected for this study will be anonymised and stored securely on Onedrive of the researcher's university, which is a secure university online repository with log ins which are controlled by the team of this research project. This repository allows us to retrieve the data at any time, but prevents unauthorised access.

What will happen to the results of the study?

The results from this study will be used in the following ways: related publications, and conference papers. If you would like to receive a summary of the findings, then please indicate this on the consent form or contact the lead researcher.

Who has reviewed the study?

This study has been reviewed and approved by the University of Northampton Research Ethics Committee on [date].

Contact for further information

If you have any questions about this study or your possible involvement then please contact me using the contact details below.

Qiuling Bi, the researcher, (Qiuling. Bi@northampton.ac.uk)

Dr. Qian Zhang, PhD Student Supervisor in the Faculty of Education and Humanities, (Qian.Zhang@northampton.ac.uk)

Thank you for considering taking part in this study.

Appendix K3. Participant Information Sheet for CT tests

Participant Information Sheet (CT tests)

Research Title

Critical Thinking Instruction Through Project-based Learning in Chinese EFL classes- A Case Study in HE

Why have I been invited?

You are being invited to start to take part in this research study on 02/09/2019. Before you decide whether you wish to participate, it is important for you to understand why the study is being conducted and what it will involve. Please take some time to read the information provided and discuss it with others if you wish. Please ask if there is anything that is not clear, or if you would like more information.

What is the purpose of the study?

The aims of this study are to investigate changes in your critical thinking ability, dispositions and writing performance.

Why have I been chosen?

You have been asked to take part in the study because you will be undertaken the program of College English in which some innovative approaches are included. Your response to the survey will enable this approaches to be evaluated.

Do I have to take part?

Taking part is entirely voluntary. If you decide to take part, you will n to sign a consent form to confirm that you understand the project and are willing to participate. If you decide to withdraw, you can do that before I make the data anonymous.

What will my participation involve?

Once you have agreed to take part, you will be asked to take part in three tests before and after project-based learning class to investigate the improvement of your critical thinking and writing performance. The first test, California Critical Thinking Skills Test (CCTST) is to test your critical thinking skills. The second one, California Critical Thinking Disposition Inventory (CCTDI) is to test your critical dispositions

and the final one, writing part in College English Test-Band 4 (CET-4) is to test your writing abilities.

What are the possible benefits of taking part?

The information obtained from this study will be used to complete this research project, and then will be publicised through journal articles.

What are the possible risks or disadvantages of taking part?

We do not believe that there are any possible risks. However, we will monitor data as it is collected and discuss implications of any risk which emerge with the participants involved.

What if something goes wrong?

If you have any concerns about any aspect of the way you have been approached or treated during the course of this study, then please contact the researcher at the University of Northampton, Qiuling Bi (Qiuling. Bi@northampton.ac.uk).

Will my information be kept confidential?

All the information collected for this study will be anonymised and stored securely on Onedrive of the researcher's university, which is a secure university online repository with log ins which are controlled by the team of this research project. This repository allows us to retrieve the data at any time, but prevents unauthorised access.

What will happen to the results of the study?

The results from this study will be used in the following ways: related publications, and conference papers. If you would like to receive a summary of the findings, then please indicate this on the consent form or contact the lead researcher.

Who has reviewed the study?

This study has been reviewed and approved by the University of Northampton Research Ethics Committee on [date].

Contact for further information

If you have any questions about this study or your possible involvement then please contact me using the contact details below.

Qiuling Bi, the researcher, (Qiuling. Bi@northampton.ac.uk)

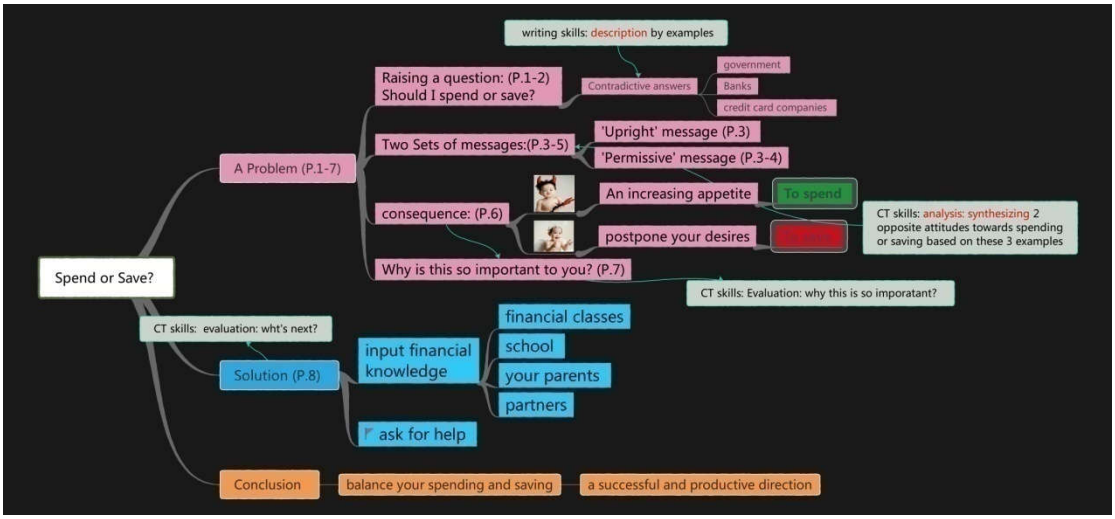
Dr. Qian Zhang, PhD Student Supervisor in the Faculty of Education and Humanities, (Qian.Zhang@northampton.ac.uk)

Thank you for considering taking part in this study.

Appendix L. Two Example Thinking Maps

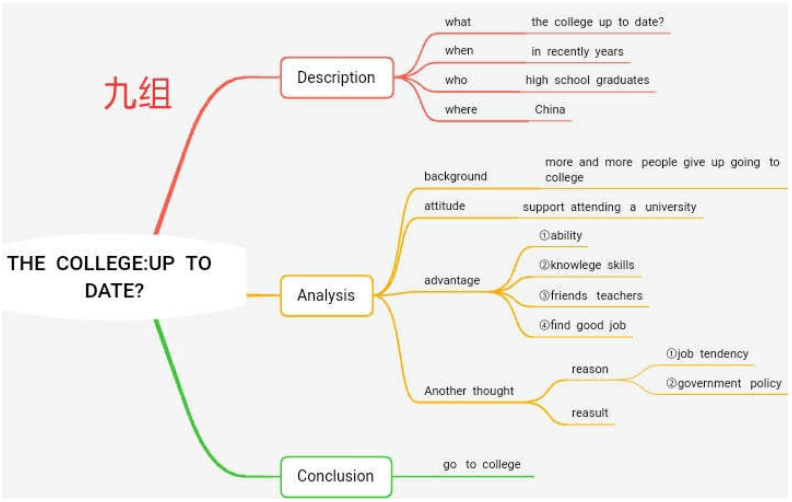
L1. A Teacher's Example Thinking Map

(*New Horizon College English*, Book 2, Unit 4: Spend or Save)

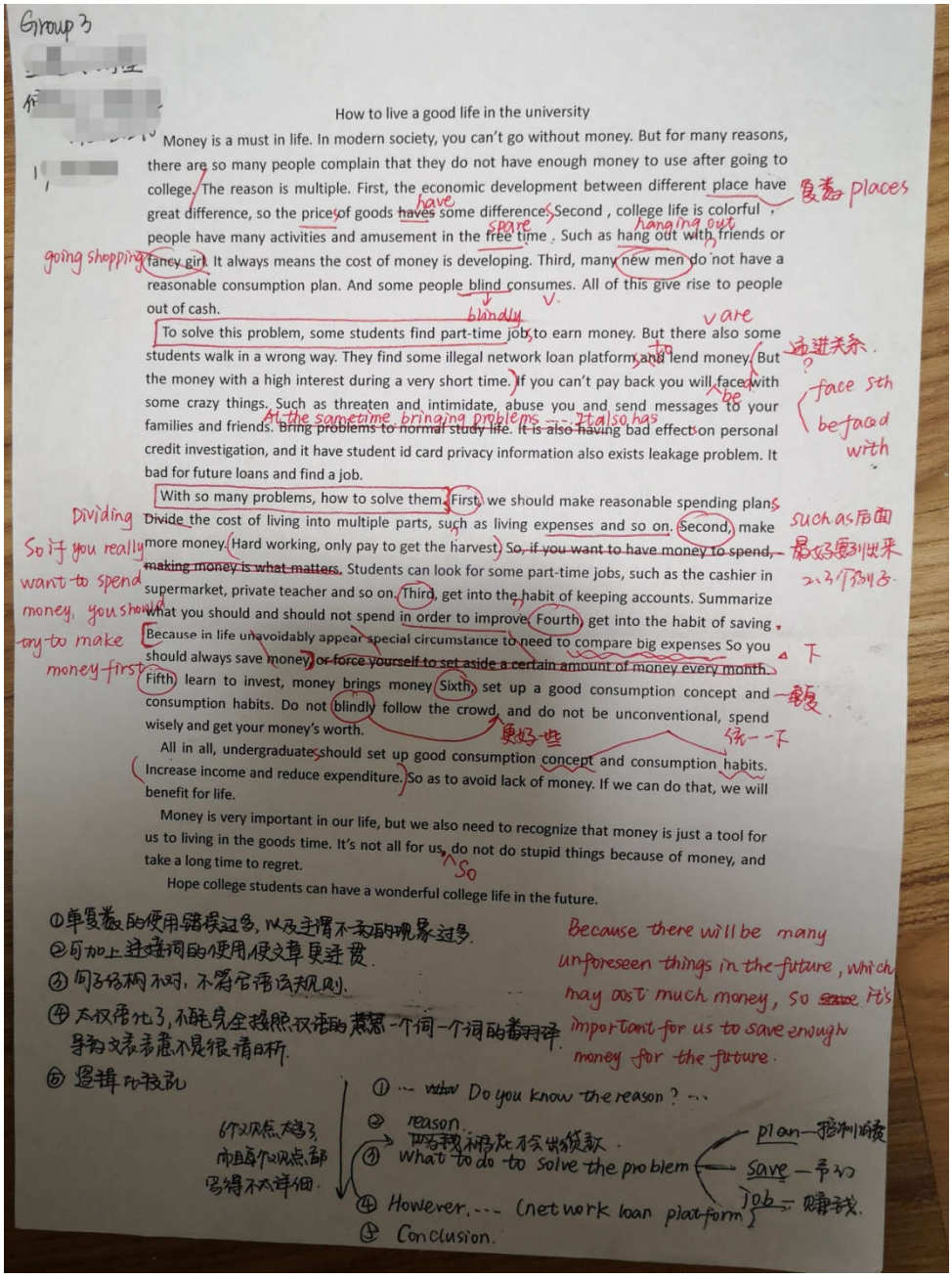


L2. A student's Example Thinking Map

(Group 9, Unit 1: College Education: Out of Date?)



Appendix M. An Example Peer Review Essay



Group 3

How to live a good life in the university

Money is a must in life. In modern society, you can't go without money. But for many reasons, there are so many people complain that they do not have enough money to use after going to college. The reason is multiple. First, the economic development between different place have great difference, so the price of goods haves some difference. Second , college life is colorful , people have many activities and amusement in the free time . Such as hang out with friends or fancy girl. It always means the cost of money is developing. Third, many new men do not have a reasonable consumption plan. And some people blind consumes. All of this give rise to people out of cash.

To solve this problem, some students find part-time job to earn money. But there also some students walk in a wrong way. They find some illegal network loan platform and lend money. But the money with a high interest during a very short time. If you can't pay back you will face with some crazy things. Such as threaten and intimidate, abuse you and send messages to your families and friends. Bring problems to normal study life. It is also having bad effect on personal credit investigation, and it have student id card privacy information also exists leakage problem. It bad for future loans and find a job.

With so many problems, how to solve them. First, we should make reasonable spending plans Divide the cost of living into multiple parts, such as living expenses and so on. Second, make more money. Hard working, only pay to get the harvest. So, if you want to have money to spend, making money is what matters. Students can look for some part-time jobs, such as the cashier in supermarket, private teacher and so on. Third, get into the habit of keeping accounts. Summarize what you should and should not spend in order to improve. Fourth, get into the habit of saving. Because in life unavoidably appear special circumstance to need to compare big expenses So you should always save money, or force yourself to set aside a certain amount of money every month. Fifth, learn to invest, money brings money. Sixth, set up a good consumption concept and consumption habits. Do not blindly follow the crowd, and do not be unconventional, spend wisely and get your money's worth.

All in all, undergraduates should set up good consumption concept and consumption habits. Increase income and reduce expenditure. So as to avoid lack of money. If we can do that, we will benefit for life.

Money is very important in our life, but we also need to recognize that money is just a tool for us to living in the goods time. It's not all for us, do not do stupid things because of money, and take a long time to regret.

Hope college students can have a wonderful college life in the future.

Because there will be many unforeseen things in the future, which may cost much money, so save it's important for us to save enough money for the future.

Appendix N. Two Example Essays before and after PBL Teaching Intervention

(Unit 5: Spend or Save)

Model 1 (An essay written by S28 before PBL teaching)

writing title: The management of the money

Nowadays, there is a problem that the undergraduates have to face is the management of the money. Do you have such a question for this?

China today's children lack of reasonable and moderate consumption idea. For the children and willing to spend money, is the modern young parents' love. The love of parents, with the rapid development of commodity economy, will push the child prematurely before money. Currently, parents treat their children on the consumption idea of misunderstanding between the two: one is to let children blindly motivated by money, "stop poor poor children" is what they advocate the slogan. In this way, can make the children grew up in a pile of money, excessive emphasis on material comforts, vain, lack of the spirit of arduous struggle and perseverance. One is that money is not a good thing, and it is not a good thing for children to have money.We should still have a clear direction for our expenses, so accounting is a very important thing. For different students, the direction of spending is different, girls are more inclined to clothes and cosmetics, while boys are more inclined to games and sports. When consumption must choose useful things for themselves, should not waste a lot of money on some useless things.

What we should do is not only plan to spend money, but also learn how to save money.

The past ten years have seen a dramatic increase in the number of students spending a lot of money on college. This phenomenon will bring heavy burden to the family, so how to save money has become an important issue. There are many ways to save money. First, we can reduce the unavoidable expenses, they can help us save more money. Second, we can do some part-time jobs and earn some money. Finally, we can also make some investments that will help us receive the money at the time of shooting. In my opinion, as a college student, we must save money very seriously. It can not only help the family light burden, but also can help us into a good habit.

We should exercise our ability to manage capital as early as possible, so as to improve our ability to survive, so that we can better adapt to the life in the society alone. Learning to make reasonable use of limited funds is also a kind of exercise for oneself. Especially at the present stage, learning to make reasonable use of living expenses can help us to have a better college life, and can also avoid constantly asking parents for money and increasing their burden.

As an adult, this is a basic ability we must have, we should learn to spend money reasonably, to avoid the problem of spending more than income. In the case of

insufficient living expenses, we can also choose to use our spare time to do part-time jobs to increase our income. It's also important to learn how to save money so we don't get caught up in the tricky business of spending money. Therefore, we should cultivate our capital management ability as early as possible, so as to prepare ourselves for the society in the future.

Model 2 (An essay written by S28 after PBL teaching)

The management of money

Nowadays, the fact that undergraduates have to face the trouble of money management has raised a prevailing topic. Owing to the fact that lack of reasonable and moderate management idea, a host of Chinese teenagers are used to waste cash unintentionally. They always complain that living costs is accountable, nevertheless, they never endeavor to control themselves even if just for a while. Do you have a problem like this?

Initially, keeping accounts is a helpful skill for money management. Some students choose to write down what they did and which was the largest part of their expenditures, making them know themselves pretty well. Nonetheless, another group of people, just spend money when they are willing to. Accordingly, the only thing they know is the remaining sum of their credit card. To be honest, I used to spend plenty of money on many useless things just because of my impulsion. The moment I realized that things became really serious is that one day I opened my wardrobe feeling nothing to wear but seeing terrible mess filling my chest up. What did I do ?Ever since then , I have thought twice before making decisions to buy something.

To figure out the reason of such kind of the schema of Chinese teenagers, it is not hard to see the correlation between it and the family education. Currently, Chinese parents treating their children on the consumption idea have two misunderstandings. One is that having children motivated by money blindly , "Stop poor poor children" is what they advocate the slogan. This way can make their children grow up in a pile of money, excessive emphasis on material comforts and lack of the spirit of arduous and perseverance. Additionally, another is that they told their children money is not a desirable thing and it is bad to have money for children. This is a wrong way, either. Simply because it is true that we cannot live without money.

The past ten years have seen a dramatic increase of the number of the students spending in the college, which means heavy burden hitting the family. So how to save money has become an significant issue. There are a variety of ways to save money.

Initially, why not reduce the unavoidable expenses?It enables us to save more money. What is more, undertaking various part-time jobs and earning some money sound like excellent ideas. Eventually, we might also make some investments which will help us to receive the money at the time of shooting. As far as I'm concerned, as a college student, we must save money seriously because it can not only help the family alleviate burden, but also help us establish a clear and correct conception for our expenses and develop a good habit.

We are supposed to improve our ability of manage capital as early as possible so that we can improve our ability to survive, thereby adapting to the life in the society alone better. Learning to make reasonable management of limited funds is also a kind of exercise for oneself. Especially at the present stage, it can help us to have a better college life, also will avoid constantly asking parents for money and increasing their burden.