

The University of Canterbury

An Investigation of the Use of Cooperative Learning

in Teaching English as a Foreign Language

with Tertiary Education Learners in China

A thesis submitted in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy in Education
in the University of Canterbury

By

Huiping Ning

November 2010

Contents

Acknowledgments	i
Dedication	ii
Abstract	iii
List of Tables	v
List of Figures	vi
List of Appendices	viii
List of Abbreviations	ix
CHAPTER ONE: INTRODUCTON	1
Overview	1
Importance of English Language Teaching in China	1
The Traditional English Language Teaching Approach in China: Reasons and Consequences	3
College English Teaching in China	7
<i>Characteristics of College English Teaching</i>	8
<i>National College English Reform</i>	9
Cooperative Learning	12
Research Questions	15
CHAPTER TWO: LITERATURE REVIEW	16
Overview	16
History of Cooperative Learning	16
Theoretical Roots of Cooperative Learning	18
<i>Social Interdependence Theory</i>	18
<i>Cognitive Development Theory</i>	19
<i>Motivational Theory</i>	20

<i>Behavioural Learning Theory</i>	22
<i>Humanistic Psychology</i>	23
Types of Cooperative Learning Methods	24
<i>Learning Together</i>	25
<i>Student Team Learning</i>	26
<i>The Structural Approach</i>	27
<i>Jigsaw</i>	29
<i>Group Investigation</i>	30
<i>Complex Instruction</i>	32
Basic Principles and Elements of Cooperative Learning	33
Positive Outcomes of Cooperative Learning	35
Key Issues in Implementing Cooperative Learning	37
<i>Grouping Students</i>	37
<i>Functioning as a Cooperative Learning Group</i>	43
<i>Selecting Cooperative Learning Methods and Techniques</i>	45
<i>Assessing Cooperative Learning Group Work</i>	47
<i>Teacher's Roles in Cooperative Learning</i>	51
Cultural Appropriateness of Cooperative Learning in China	53
Links between Cooperative Learning and Second Language/Foreign Language Teaching	58
<i>Input Hypothesis</i>	58
<i>Interaction Hypothesis</i>	59
<i>Output Hypothesis</i>	60
<i>Affective Factors</i>	61
<i>Goals of Second Language/Foreign Language Teaching</i>	63
Links between Cooperative Learning and Tertiary Students	66
Links between Cooperative Learning and Large-Class Teaching	69
Recent Studies on Using Cooperative Learning in Teaching English as a Second Language/Foreign Language	71

<i>Experimental Studies in Different Countries</i>	71
<i>Recent Studies in China</i>	74
 CHAPTER THREE: METHODOLOGY	 77
Overview	77
Use of Quantitative Research Methodology	77
Participants	79
Research Procedure	80
<i>Pilot Study</i>	81
<i>Main Study</i>	86
Intervention Procedure	87
<i>The Adapted Cooperative Learning Method Used with the Intervention Group</i>	88
<i>Traditional Instruction Used with the Comparison Group</i>	96
<i>Summary of Contrasts between Cooperative Learning and Traditional Teaching</i>	97
Measures	98
<i>College English Test</i>	99
<i>Language Learning Orientations Scale</i>	100
<i>Social Skills Scale for Chinese College English Learners</i>	102
Data Analysis	109
 CHAPTER FOUR: RESULTS	 113
Overview	113
Results from the College English Test	113
Results from the Language Learning Orientations Scale	121
Results from the Social Skills Scale for Chinese College English Learners	130
<i>Results on the Eight Subscales and Total Scores</i>	130
<i>Results from Items Focusing on Learning of English</i>	141

CHAPTER FIVE: DISCUSSION	154
Overview	154
English Language Proficiency	154
Motivation for Learning English	160
Social Skills	167
English-Learning-Related Social Skills	171
Implications for Practice	177
<i>Design of Appropriate Cooperative Learning Tasks</i>	177
<i>Extra Workload Involved in Preparing and Implementing Cooperative Learning Lessons</i>	
<i>Lessons</i>	180
<i>Limited Teaching Hours and a Large Curriculum to Cover</i>	185
<i>Students' Use of the First Language in Teamwork</i>	188
Contributions of this Research	189
Limitations of this Research	190
Recommendations for Future Research	194
REFERENCES	196
APPENDICES	213

Acknowledgements

I would like to express my gratitude to all the people who have contributed to the completion of this thesis:

Professor Garry Hornby, for his excellent supervision and careful guidance throughout this long and tough journey;

Dr. Fengyuan Ji, Marcia Pilgrim, Dr. Yaping Chen, for their kind support and helpful feedback on my research;

Dr. Pat Coope and Prof. John Everatt for their advice on statistics;

Dr. Kimberly Noels, for her permission regarding the use of the Language Learning Orientations Scale developed by her;

Pat Syme, for her professional advice, careful reading of every draft of my thesis, continued encouragement and assistance;

and

Wei Zhuang, my beloved husband, for his full support all the way.

Dedication

This thesis is dedicated to the memory of my grandma, Goulan Zhang

Abstract

This thesis adapts cooperative learning methods for the College English teaching context in China. Its focus is on investigating the effects of cooperative learning on students' English language proficiency, learning motivation and social skills, in comparison with traditional whole-class instruction, by employing a pre-test-post-test control group quasi-experimental design.

The first chapter clarifies the context for the research, which includes an introduction about the importance of English language teaching in China, a description of the widely used traditional approach, as well as its negative consequences. Observations are made regarding the characteristics of College English teaching and recent nationwide College English reform, which have spurred a transformation of the traditional approach at the tertiary level, with a focus on enhancing students' listening and speaking abilities in English. This is followed by a brief overview of cooperative learning, as well as its potential to contribute to College English teaching. The overall aim of the research and the specific research questions addressed are presented at the end of this chapter.

The second chapter consists of a review of the literature regarding the history of cooperative learning, the fundamental theoretical underpinnings of cooperative learning, major cooperative learning methods, basic elements of cooperative learning, and its positive outcomes. This chapter also elaborates on key issues in implementing cooperative learning in the classroom and the cultural appropriateness of cooperative learning in China. In accordance with the specific context for this research, this chapter addresses the relevant connections of cooperative learning to second and foreign language teaching, tertiary education, and large-class instruction. An extensive review is also included regarding recent studies on using cooperative learning in second and foreign language teaching around the world as well as in China. The review locates a gap in the existing studies—the effectiveness of cooperative learning in teaching English as a foreign language to tertiary learners in China—which constitutes the focus of this thesis.

The third chapter starts with some key concepts essential for quantitative methodology used in

this research. It is followed by an introduction of participants and the general research procedure, which includes a pilot study and a main study employing a pre-test-post-test control group quasi-experimental design. Details of the intervention procedure are provided, focusing on different teaching methods used in the cooperative learning classroom and the traditional classroom. This chapter also provides details of the three measures used in this research: the College English Test, the Language Learning Orientations Scale, and the Social Skills Scale for Chinese College English Learners. At the end of the chapter there is an explanation of specific techniques and principles for data analysis.

In chapter four, results are presented based on analysis of the data from the three measures. In general, the results focus on seven aspects: mean scores, standard deviations of pre- and post-test scores for each group, effect sizes of Cohen's *d* from pre- to post-test for each group, alpha values of paired-samples *t*-tests for each group, alpha values of interaction effect between group and time from ANOVAs, alpha values of post-test difference between groups from one way ANCOVAs, and means plots for each of testing areas. The findings provide evidence in favour of cooperative learning in some areas, for instance, in teaching speaking, listening and reading, in generating intrinsic motivation, and in incorporating students' equal participation and individual accountability into learning.

The final chapter includes a discussion of the findings on the three measures in relation to the findings of previous research. It goes on to discuss implications for the practice of English language teaching, with a focus on the challenges of using cooperative learning in Chinese tertiary institutions. These challenges mainly include designing appropriate cooperative learning tasks, extra workload involved in preparing and implementing cooperative learning lessons, limited teaching hours and a large curriculum to cover, as well as students' use of the first language in teamwork. The chapter ends with a discussion about the major contributions and limitations of the current study, as well as recommendations for future research.

List of Tables

Table 1: Demographic information on participants in the three groups	80
Table 2: Summary of results on the CET for the intervention and comparison groups	114
Table 3: Summary of results on the LLOS for the intervention and comparison groups	122
Table 4: Summary of results on the SSS-CCEL for the intervention and comparison groups	131
Table 5: Summary of results from items focusing on learning of English for the intervention and comparison groups	142

List of Figures

Figure 1: Components and procedures of an adapted CL method with Chinese EFL learners	89
Figure 2: Pre- and post-test mean listening scores on the CET	115
Figure 3: Pre- and post-test mean speaking scores on the CET	116
Figure 4: Pre- and post-test mean reading scores on the CET	117
Figure 5: Pre- and post-test mean writing scores on the CET	118
Figure 6: Pre- and post-test mean vocabulary scores on the CET	119
Figure 7: Pre- and post-test mean scores of the total on the CET	120
Figure 8: Pre- and post-test mean scores of intrinsic motivation on the LLOS	123
Figure 9: Pre- and post-test mean scores of integrated motivation on the LLOS	124
Figure 10: Pre- and post-test mean scores of identified motivation on the LLOS	125
Figure 11: Pre- and post-test mean scores of introjected motivation on the LLOS	126
Figure 12: Pre- and post-test mean scores of external motivation on the LLOS	127
Figure 13: Pre- and post-test mean scores of amotivation on the LLOS	128
Figure 14: Pre- and post-test mean scores of the total on the LLOS	129
Figure 15: Pre- and post-test mean scores of self-confidence on the SSS-CCEL	132
Figure 16: Pre- and post-test mean scores of sense of cohesion on the SSS-CCEL	133
Figure 17: Pre- and post-test mean scores of initiative in socialization on the SSS-CCEL	134
Figure 18: Pre- and post-test mean scores of being positive on the SSS-CCEL	135
Figure 19: Pre- and post-test mean scores of checking for understanding on the SSS-CCEL	136
Figure 20: Pre- and post-test mean scores of equal participation and accountability on the SSS-CCEL	137
Figure 21: Pre- and post-test mean scores of acceptance and empathy on the SSS-CCEL	138
Figure 22: Pre- and post-test mean scores of conflict management on the SSS-CCEL	139
Figure 23: Pre- and post-test mean scores of the total on the SSS-CCEL	140
Figure 24: Pre- and post-test mean scores of item 3 for focus skills on the SSS-CCEL	143

Figure 25: Pre- and post-test mean scores of item 4 for focus skills on the SSS-CCEL	144
Figure 26: Pre- and post-test mean scores of item 7 for focus skills on the SSS-CCEL	145
Figure 27: Pre- and post-test mean scores of item 10 for focus skills on the SSS-CCEL	146
Figure 28: Pre- and post-test mean scores of item 15 for focus skills on the SSS-CCEL	147
Figure 29: Pre- and post-test mean scores of item 25 for focus skills on the SSS-CCEL	148
Figure 30: Pre- and post-test mean scores of item 26 for focus skills on the SSS-CCEL	149
Figure 31: Pre- and post-test mean scores of item 27 for focus skills on the SSS-CCEL	150
Figure 32: Pre- and post-test mean scores of item 35 for focus skills on the SSS-CCEL	151
Figure 33: Pre- and post-test mean scores of the total for focus skills on the SSS-CCEL	152

List of Appendices

Appendix 1: A summary of teaching content of College English course for reading and writing covered in the intervention	213
Appendix 2: A summary of teaching content of College English course for listening and speaking covered in the intervention	215
Appendix 3: A sample assignment of a team-task involving elements of Jigsaw, Complex Instruction and Group Investigation	216
Appendix 4: An example of a worksheet for teaching reading	218
Appendix 5: A team-grade sheet used by teams for peer- and self-grading	219
Appendix 6: A teamwork recording sheet for the teacher's use	220
Appendix 7: Contrasts of a typical 50-minute session between the two different methods for teaching listening and speaking	221
Appendix 8: Contrasts between the two different methods for teaching writing	224
Appendix 9: Contrasts between the two different methods for teaching vocabulary	225
Appendix 10: A sample of College English Test paper	227
Appendix 11: Scoring criteria for speaking on the CET	236
Appendix 12: The Language Learning Orientations Scale in both English and Chinese Versions	238
Appendix 13: The Social Skills Scale for Chinese College English Learners in both English and Chinese versions	242
Appendix 14: Between-group comparisons of ages, years of learning English, and pre-test scores in 19 areas on the three measures of the CET, the LLOS, and the SSS-CCEL	246

List of Abbreviations

1. CET for College English Test
2. CHC for Confucian Heritage Culture
3. CL for cooperative learning
4. EFL for English as a foreign language
5. ELT for English language teaching
6. FL for foreign language
7. GI for Group Investigation
8. LLOS for Language Learning Orientations Scale
9. L2 for second language
10. NHT for Numbered-Heads-Together
11. SSS-CCEL for Social Skills Scale for Chinese College English Learners
12. STAD for Student-Team Achievement-Division
13. STT for student talk time

CHAPTER ONE: INTRODUCTION

Overview

This chapter is aimed at clarifying the context for this research. It starts with an introduction about the importance of English language teaching in China. Then it moves to a description of the widely used traditional approach, which is considered to be of limited effectiveness in developing students' communicative competence. This is followed by a discussion of the reasons for and negative consequences of using this approach. Recent nationwide College English teaching reform in China has spurred the transformation of the traditional approach at the tertiary level, with the implementation of new curriculum requirements which emphasize the development of students' listening and speaking abilities in English. This chapter clarifies the situation of College English teaching within this nationwide reform. It is suggested that cooperative learning (CL) has the potential to contribute to College English teaching, meet the new curriculum requirements, and enhance a wide range of learning outcomes. The overall aim of the research and the specific research questions addressed are presented at the end of this chapter.

Importance of English Language Teaching in China

English has become widely accepted as a global language because of its critical communication role in the field of business, technology, science, the Internet, popular entertainment and even sports (Nunan, 2003b). The global status of English has inevitably attracted attention to the quality of English language teaching (ELT) all round the world. China, with its rapid growth in socioeconomic developments and its important roles in many international organizations, is no exception due to its urgent need for a large number of competent English language users who can

interact within a global society. At the national level, “English has been progressively linked to China’s open-door policy, the modernisation and rapid development of the economic systems, China’s burgeoning international trade and the recognition of China’s significant role in world affairs” (Jin & Cortazzi, 2004, p. 119). Additionally, recent events, such as hosting the 2008 Olympic Games and the 2010 World Expo, have acted as a kind of catalyst for expanding the scale of and enthusiasm for learning English. For instance, to be a successful host of the 2008 Olympic Games, the Chinese government once announced that the whole nation—not only students but also people of all ages and occupations—should learn English, so that visitors to China could be warmly welcomed in this global language (Guo, 2001). Nowadays in China, English learning takes place at all levels of the education system, with a total of about 300 million English learners (Jin & Cortazzi, 2004; Yang & Nicholas, 2008), and this number is increasing all the time.

Formal schooling in China is generally composed of three levels: six years of elementary education in primary school for students aged from seven to 12 years; six years of secondary education which includes three years in the junior middle school and three years in the senior middle school; and higher education which typically includes four years of undergraduate studies and three or more years of postgraduate studies. For about 30 years, English has been taught as a formal requirement from secondary school to tertiary level. Since September 2001, a policy from the Ministry of Education has lowered the age at which compulsory English instruction begins, from 11 to nine years of age. That is, English has become a compulsory course from Grade Three in primary school through to postgraduate studies. Typically, students take regular English courses four sessions a week, 16-18 weeks a term, for eight terms in primary school, 12 terms in middle school, and four to eight terms at university (Wu, 2001). These numbers reflect the massive investment of the Chinese government in ELT and the extra emphasis on English courses in formal education. In China, there is a widespread perception that a good command of English confers prestige on individuals and opens doors to academic and professional success (Jin & Cortazzi, 2004). This perception spurs students and their parents to invest considerable time, energy and money in the learning of English.

A decisive factor in the importance of English is linked to the key role of English in a variety

of gate-keeping and high-stakes examinations, two of which are the senior middle school entrance examination and the national college entrance examination, administered annually to middle-school students. Due to nationwide nine-year compulsory education, every child is eligible to receive education in primary and junior middle schools. But their opportunities for higher level education are decided by how well they do on the relevant competitive entrance examinations, in which English constitutes a major component of testing. An admission to a good senior middle school is what students and their parents spare no effort to achieve, since this is considered the first step to getting into a good university. In the national college entrance examination “the role of English was significantly increased from 100 points to 120 points in the early 1990s and more recently to 150 points. Other subjects, except Chinese, received 100 points or less” (Jin & Cortazzi, 2004, p. 123). According to data collected in 2004 (Shu, 2004, p. 19), 50 million enrolments in junior middle schools were reduced to 14 million at the tertiary level after the two rounds of large-scale screening through the entrance examinations. “The limited number of places in senior middle schools and higher education, even with the recent large-scale expansion, means that the competitive entrance examinations retain a strong role” (Jin & Cortazzi, 2004, p. 125). In addition, English tests are not only used with students but also with professionals in a variety of occupations (e.g. engineers, teaching staff, lawyers, accountants, economists, and doctors), who want to gain a higher professional position (e.g. from associate professor to full professor). Therefore, English has become a vital element that determines people’s education opportunities and professional success in China.

The Traditional English Language Teaching Approach in China:

Reasons and Consequences

With a general consensus that “the ability to communicate effectively in English is now a well-established goal in ELT)” (Hedge, 2000, p. 44), the critical role of the communicative

approach in ELT has been well recognized on a global scale (Bjorning-Gyde, Doogan & East, 2008; Ellis, 1994, 1999; Harmer, 1998, 2007; Hedge, 2000; Nunan, 2003a; Richards & Rodgers, 2001; Siemon, 2010; Teng, Niu & Wolff, 2004). The communicative approach emphasizes the use of authentic communicative activities and group work in the classroom, as well as the involvement of students' active participation in language learning. It is aimed at developing students' ability to use the language in real-life communication rather than merely teaching them linguistic knowledge. In the communicative classroom, a primary role of the teacher in the classroom is to be a language facilitator and help communicative groups function well. A powerful message derived from the communicative approach is that students must learn to speak by speaking. However, this perception is not reflected in the traditional ELT setting in China, where whole-class instruction and teacher talk dominate English classrooms in primary schools, middle schools and universities (Shu, 2004). "Most research indicates that Chinese students' English learning strategies are primarily focused on reading and writing, on grammar and translation, and on memorisation of vocabulary" (Siemon, 2010, p. 40).

ELT at the secondary level is typically test-oriented and teacher-centred, due to the high pressure of the two high-stakes gate-keeping entrance examinations, which focus on grammatical knowledge, accurate use of vocabulary and reading comprehension, but ignore the actual use of English. This may largely account for the continued popularity of test-oriented teaching in middle schools. For the purpose of passing these examinations, about 90 percent of students spend the whole of their last year in junior and senior middle schools preparing for the tests through repeatedly practising potential test items and doing mock test papers (Shu, 2004). As a result, the grammar-translation method, involving little attention to real-life communication, is undoubtedly considered the most convenient and possibly effective method to deal with test-related language knowledge. This teaching method will probably continue to dominate English language classrooms in middle schools while these gate-keeping tests exist. In other words, applying the communicative approach to the ELT setting at the secondary level is currently likely to be a very difficult mission.

At the tertiary level, although students are no longer under the pressure of gate-keeping

English examinations, many researchers and practitioners maintain that, typically, rote memorising, heavy grammar instruction, and detailed vocabulary and text explanation continue to characterize the teaching methods used at universities, which therefore exhibit no substantial difference from the methods prevailing in middle schools (Bjorning-Gyde et al., 2008; Hu, 2005; Jin & Cortazzi, 2004; Peng & Qi, 2006; Shu, 2004; Teng et al., 2004; Wang, 1999). A recent survey on classroom teaching methods with 40 randomly selected English teachers in a university in Jiangsu (a province, which is more economically developed and educationally advanced than the average) has found that over 70 percent of teachers still predominantly use the grammar-translation method in their classrooms (Shu, 2004). In another similar survey, Xia (2002) found that 90 percent of English teacher participants reported that their primary role in classroom teaching remained as language model and language explainer. Interestingly, 77 percent of participants reported that this traditional way of teaching failed to develop an adequate level of communicative competence on the part of students, and that they believed that an English teacher should facilitate students' active participation and actual use of the target language. Obviously, there was a mismatch between their belief about good teaching and their actual practice in the classroom; that is, they did not teach according to what they believed was good teaching.

Some researchers have investigated the reasons behind this reluctance to use the communicative approach in spite of teachers' awareness that traditional instruction is inadequate in facilitating students' actual use of English (Chen, 2007; Hu, 2002; Hu, 2005; Peng, 2007; Shu, 2004; Siemon, 2010; Wang, 1999; Xia, 2002; Yang & Nicholas, 2008). The reasons specified in these studies can be synthesized into three categories: teacher-related, student-related and situational-constraint-related. First, teachers feel comfortable when delivering teaching content by a familiar method. They are very likely to teach in the way that they were taught as students. Teachers' lack of relevant professional competence in managing group activities or their inadequate linguistic skills in authentic communication may also form barriers for those who feel like trying some pedagogical innovations grounded in the communicative approach.

The second category of reasons is related to students, who are found to be short of vocabulary to express ideas and make too many grammatical mistakes in their speaking. Based on the

secondary-level English syllabus, senior high school students are expected to master at least 2,000 English words and a wide range of grammar rules, some of which are already quite complex (e.g. use of subjunctive mood and past perfect continuous tense). At the tertiary level, students are expected to increase their vocabulary by approximately 1,500 words each year during their first two years, and more complex grammatical knowledge is introduced to them through textbooks. These are supposed to be adequate for daily communication in English. So the situation may be that “Chinese students memorise large amounts of vocabulary and grammar rules, but prove not to be proficient in using the vocabulary appropriately or applying the grammar rules” in real communication (Siemon, 2010, p. 4). In view of this situation, the problem can be interpreted in a different and probably more accurate way: that is, what students lack is active vocabulary and grammatical competence (i.e. the ability to transform grammar knowledge they have learned from textbooks into real communication). Actually, the most effective way to transfer rote-memorized vocabulary and grammatical knowledge into active vocabulary and grammatical competence is through a large amount of practice in authentic communicative settings. However, students’ inadequacy in communication cannot become an excuse for shying away from the communicative approach; but on the contrary, they should be provided with more opportunities to access authentic communication in the target language. Other common problems with students include their lack of motivation and cooperative skills for working in groups, and their reluctance to communicate with others in English.

The third category is related to some situational constraints commonly existing in tertiary ELT in China, which entail the use of textbooks, limited teaching hours and large class sizes. The text-books used at the tertiary level, especially for reading and writing, are largely designed for whole-class instruction instead of communicative teaching. Teachers have also found that communicative group work is comparatively time-consuming and not so efficient in delivering curriculum content, especially considering limited class time of typically four hours a week. Large class sizes of over 50 students constitute another barrier to the application of the communicative approach, for teachers have found it hard to effectively monitor and facilitate students’ communicative activities in large classes.

In spite of the above difficulties, transforming the traditional teaching approach has become essential due to complaints about unsatisfactory learning outcomes (Jin & Cortazzi, 2004; Li, 1996; Liu & Dai, 2004; Ruan & Jacob, 2009; Shu, 2004; Siemon, 2010; Teng et al., 2004; Wang, 1999; Yang & Nicholas, 2008). It is very disappointing that “Chinese students’ English is often referred to as ‘deaf and dumb English’” because of their poor listening and speaking abilities, after many years of English learning focusing on grammar, vocabulary and intensive reading (Siemon, 2010, p. 41). This was also reflected in a speech by Lanqing Li, the former vice premier in charge of education, at a national education conference. According to Li’s (1996) comments, ELT in China has been extremely “time-consuming and low-efficient”, and it is frustrating to see enormous investments result in such a minor achievement regarding students’ communicative competence. Although these comments may not be agreed on by everyone, the reality that many university graduates are unable to carry on simple conversations in English cannot be ignored. As for the causes of this common problem, the effectiveness of the current teaching approach has been questioned (Hu, 2005; Liao, 2004; Jin & Cortazzi, 2004; Peng, 2007; Shu, 2004; Siemon, 2010; Teng et al., 2004; Ruan & Jacob, 2009; Yang & Nicholas, 2008). Additionally, it should be noted that, compared with ELT at the secondary level, tertiary ELT has some advantages conducive to transforming the traditional teaching approach. These advantages include: tertiary learners being largely free of the pressure of gate-keeping English examinations; communicative competence being particularly emphasized by an updated guideline for tertiary ELT; and nation-wide reform having been implemented in English teaching and testing at the tertiary level with more weight placed on listening and speaking ability (these latter two points will be addressed in the next section). Therefore, changing the traditional teaching approach has become not only possible but also essential for ELT at the tertiary level in China.

College English Teaching in China

Tertiary ELT for undergraduates in China is composed of two parts—for English-majors and

non-English majors—using different curricula, course requirements and sets of textbooks. ELT for non-English majors, who make up the overwhelming majority of tertiary learners, is often termed as College English teaching and is the focus of this research. Since its establishment in the early 1980s, “CE (College English) has developed into a systematic and independent subject and has become one of the most important components of the Chinese higher education curriculum” (Ruan & Jacob, 2009, p. 467).

Characteristics of College English Teaching

College English is a curriculum-based compulsory course for non-English majors in the first and second year at universities in China. College English learners are young adults mostly aged between 18 and 23, who have learned English for about ten years. College English teaching is classroom-based with two sets of assigned textbooks, one for reading and writing, and the other for listening and speaking. There are usually four 50-minute sessions for College English each week, with 13-18 weeks each semester and a total of four semesters over two years. College English is typically taught in large classes of over 50 students due to the shortage of College English teachers and other relevant teaching resources. According to a survey in 2001, the ratio between College English teachers and learners was 1:130 in China, and this gap has probably widened as a result of expanded enrolment nationwide (Shu, 2004). As with all the other types of ELT in China, College English teaching is typically English as a foreign language (EFL) teaching since students have hardly any chance to use English outside the classroom. In most cases, there is no native-speaker English teacher available to College English learners.

The College English Test (CET) is an integral part of the assessment system for evaluation of the results of College English teaching and learning. The CET ranges from band one to band four for College English learners, who are expected to pass one band higher each semester. Usually, the CET Band One (CET-1) is administered as the final examination at the end of the first semester and the CET-4 is used at the end of the fourth semester. The CET-4 is administered nationwide, while the tests of other bands are usually conducted within a university or college, with the test

design following the format of the CET-4.

As for teaching methods widely used in the College English teaching context in China, Ruan and Jacob (2009) summarized these as follows:

Teaching methods are often tedious and out of date, with a distinct lack of variation in style.

The most common method used is the grammar-translation approach, by which the teacher occupies most of the class time imparting linguistic knowledge. Students listen passively and take notes on whatever the teacher says. Neither opportunities nor encouragement are provided to the student for the actual use of English... Teaching models become a kind of routine work, with the teacher explaining the meanings of the new words in the vocabulary list, analyzing grammatical structures, translating difficult sentence examples, and finally giving exercises to reinforce the linguistic knowledge (p. 470).

According to several nationwide surveys investigating College English learners' perceptions on current teaching, approximately 80 percent of respondents expressed their dissatisfaction with current College English teaching, especially its ineffectiveness in enhancing their listening and speaking ability (Ruan & Jacob, 2009; Shu, 2004). A recent study has found that many College English teachers believe that it is essential to implement innovations in classroom teaching practice in order to improve students' communicative competence (Wette & Barkhuizen, 2009). Also, a nationwide survey with employers conducted by the Ministry of Education indicated a widespread dissatisfaction with college graduates' communicative competence in English (MOE, 2001). Many employers repeatedly state that "they need competent users of English but not 'high scorers but poor users'" (Liu & Dai, 2004, p. 5). Apparently, there is an urgent call for a transformation of College English teaching in China, in order to meet the objective of language education, and the needs of students and employers, as well as society.

Nationwide College English Reform

In order to improve the quality of College English teaching and sort out existing problems, College English teaching reform has been highlighted since the early 2000s, at all levels of the

educational system from individual higher education institutions to the Ministry of Education. A very important action taken in the reform has been the updating of the old national College English syllabus which mainly focused on the development of students' reading ability. The updated syllabus, entitled "College English Curriculum Requirements" (MOE, 2004; 2007), was first implemented as a trial in 2004, and then in 2007 it was formally issued by the Ministry of Education as the nation-wide guideline for College English teaching. "The theme of the new teaching requirements can be boiled down to one expression: 'Listening and speaking are to be particularly stressed'" (Hu, 2005, p. 6). According to the new requirements, the objective of College English is as follows (MOE, 2007):

The objective of College English is to develop students' ability to use English in a well-rounded way, especially in listening and speaking, so that in their future studies and careers as well as social interactions they will be able to communicate effectively, and at the same time enhance their ability to study independently and improve their general cultural awareness so as to meet the needs of China's social development and international exchanges (p. 25).

In addition, the new requirements also advocate that computer- or web-based teaching models should form an integral part of College English teaching in order to alleviate the shortage of teaching staff and learning resources. However, some critical obstacles seriously hamper the actual implementation of this policy. In China, apart from a small number of top universities directly under the administration of the Ministry of Education, the majority of universities and colleges often suffer from inadequate financial support. Computer-supported or web-based teaching models involve enormous financial investment to provide necessary facilities (e.g. adequate number of language laboratories and self-study computer rooms), technological support (e.g. supply of suitable software in terms of learning materials, teaching management and assessment; reliable intranet or campus network; proper technological maintenance service), and relevant training programs to prepare teachers for their new roles in using advanced technology in teaching. None of these obstacles can be easily overcome, especially considering the large number of universities and College English learners in China. Therefore, it is most likely that classroom-based teaching

models will continue to play a dominating role in ELT in many universities and this situation cannot be changed within a short period of time in China. In other words, practical classroom-based innovations which facilitate the improvement of College English learners' communicative competence are essential for the current College English teaching situation in China.

In order to facilitate the implementation of the new requirements, some specific measures and recommendations are followed. First of all, the new syllabus requires that the frequency of teaching sessions in the language laboratory, which usually focus on the teaching of listening and speaking, be doubled from one hour to two hours a week. This could increase students' access to more authentic English from CD-ROMs, video recordings and other learning resources, and would undoubtedly benefit their actual use of English. Another initiative to cope with the new requirements was the reform of the CET, with the aim of developing and using more items for "the testing of integrative and practical abilities, particularly the abilities of listening and speaking" (Hu, 2005, p. 7). Additionally, it is clearly stated that the CET should serve teaching and learning instead of directing them. In order to avoid the CET dominating the assessment system and directing College English teaching and learning it is recommended that teachers include formative assessment as an integral part of the overall course evaluation for the College English course (Liu & Dai, 2004; MOE, 2007).

Therefore, transformation of the traditional teaching approach has become a prerequisite for the fulfilment of the new requirements and the success of College English reform. As for the most feasible way to transform the teaching approach, many researchers and practitioners assert that it is not workable to simply import western pedagogical philosophies and practices into the educational setting in China (Bjorning-Gyde et al., 2008; Hu, 2005; Jin & Cortazzi, 2004; Li, 2008; Liao, 2004; Peng, 2007; Shu, 2004; Siemon, 2010; Teng et al., 2004; Wang, 1999, Wu, 2001). They strongly advocate the development of an appropriate eclectic teaching model which combines traditional teacher-centred whole-class instruction with a student-centred communicative approach. In recent years, various ELT approaches and relevant innovations have been introduced and tried out in College English classrooms. These include communicative

language teaching, task-based learning, interactive learning, computer-assisted language teaching, and autonomous learning.

Faced with such a galaxy of innovations, it would be difficult to say which are truly effective. In this sense, reform of foreign language education is to encourage people to move from old practices to new philosophies generally, and to ascertain which approaches are most promising particularly. Whether you agree with a particular method or not, reform is a touchstone. Your approval or disapproval can only be answered or tested in the reform process (Hu, 2005, p. 5).

Therefore, nationwide College English teaching reform has created a favourable environment to investigate the feasibility and effect of many instructional innovations, one of which is cooperative learning.

Cooperative Learning

The importance of group work in developing students' communicative competence has been well recognized in the literature on second and foreign language teaching (Bailey, 2003; Ellis, 1994, 1999, 2009b; Harmer, 1998; 2007; Hedge, 2000; Nunan, 2003a; Richards & Rodgers, 2001). The appropriate use of group work is considered the core of many modern teaching approaches, such as communicative language teaching, task-based language learning and interactive learning. Moreover, group skills have become essential for success in more and more professional contexts, which are increasingly characterized by intensive division of labour and thus call for a wealth of collaboration and teamwork across professions and cultures (Baloche, 1998; Gillies, 2007; Kagan, 1994; Kagan & Kagan, 2009; Phipps, Phipps, Kask & Higgins, 2001). "These group skills have to be learned, however, and the classroom is a good place to begin this type of training" (Phipps et al., 2001, p. 20).

Cooperative learning (CL) is defined as the "instructional use of small groups so that students work together to maximize their own and each other's learning" (Johnson, Johnson & Holubec,

1998, p.1:5). It is often implemented through a set of “highly structured, psychologically and sociologically based techniques” (Oxford, 1997, p. 444). A key point to accurately understand CL is that not all group work constitutes CL. What makes CL differ from other types of group work largely lies in its two fundamental elements: positive interdependence and individual accountability (Baloché, 1998; Brown & Thomson, 2000; Dishon & O’Leary, 1998; Gillies, 2007; High, 1994; Holt, 1993; Hornby, 2009; Johnson et al., 1998; Kagan, 1994; Kagan & Kagan, 2009; McCafferty, Jacobs & DaSilva Iddings, 2006; Slavin, 1995). Integration of positive interdependence into group work is very likely to result in mutual support and good cooperation among team members. Positive interdependence also generates peer norms favouring achievement, increases the quantity and quality of peer interaction, and thus creates a supportive and non-stressful learning environment. When students are clear about their individual accountability and specific roles in group work, they are more likely to engage in active participation and feel motivated to learn.

The use of CL in language teaching is an extension of the principles of the communicative approach, which is defined as the “systematic and carefully planned use of group-based procedures in teaching as an alternative to teacher-fronted teaching” (Richards & Rodgers, 2001, p. 196). The goal of using CL in language teaching is to provide maximum development of communicative competence by increasing authentic peer interaction and mutual support in groups. CL group work is likely to produce a favourable language learning environment where supportive peers feel motivated and obliged to produce language output and provide comprehensible input, while feeling safe taking risks trying out the language in authentic situations (High, 1994; Holt, 1993; Jacobs & Goh, 2007; Long & Porter, 1985; McCafferty et al., 2006; Oxford, 1997; Richards & Rodgers, 2001). Thus CL is assumed to be effective in terms of providing opportunities for increased meaningful language production, and allows learners to use the language in a natural, supportive and safe environment.

Access to authentic interaction with peers in the classroom is particularly valuable for EFL language learners such as College English learners in China, who have very few chances to use English outside the classroom since English is not the language of communication in the society.

CL is also likely to provide a solution to some common problems with Chinese EFL learners in group work, who tend to lack the necessary cooperative skills as well as motivation and willingness to communicate in English with each other. Importantly, CL and traditional instruction do not always stand in contrast to each other. Some CL models integrate the two different approaches, by drawing useful elements from both whole-class instruction and communicative group work (Slavin, 1995). This facilitates a smooth transition from one approach to another, and follows the suggestion of using an eclectic teaching model which appropriately balances the use of old practices and new techniques (Bjorning-Gyde et al., 2008; Hu, 2005; Jin & Cortazzi, 2004; Li, 2008; Liao, 2004; Peng, 2007; Shu, 2004; Siemon, 2010; Teng et al., 2004; Wang, 1999, Wu, 2001). Overall, CL may serve as a powerful tool for College English teaching because it facilitates an instructional context that supports many aspects relating to language development. It has the potential to suit the College English teaching context and enhance students' learning outcomes.

The CL approach has been tried out and well researched for many decades in the West, especially as a response to the obligatory integration of public schools in the United States since 1970s. "Cooperative learning has perhaps the strongest empirical research base of any educational innovation" and there is probably more evidence supporting the use of CL than there is for any other aspect of education (Kagan & Kagan, 2009, p. 1:5). According to Johnson et al. (1998), there are many hundreds of studies which demonstrate the positive effects of CL across a wide range of subject areas and age groups, in comparison with traditional instruction stressing individualistic or competitive learning. The positive outcomes of CL are found to be related to a wide range of aspects including academic achievements, social or emotional development, cognitive growth, psychological health and learning motivation. Theoretically, CL can be applied to teach any subject matter at any educational level from kindergarten to university (for reviews, see Baloche, 1998; Hornby, 2009; Kagan, 1994; Johnson et al., 1998; Johnson, Johnson & Smith, 2007; Sharan, 1994; Slavin, 1995; Totten, Sills, Digby & Russ, 1991).

However, the impact of CL on classroom organization and procedures both with foreign language teaching (Holt, 1993; Jacobs & Goh, 2007; McCafferty et al., 2006; Oxford, 1997; Richard & Rodgers, 2001) and at the post-secondary level (Baer, 2003; Riordan, Street & Roof,

1997; Sharan, 1994; Slavin, 1997) have not been systematically studied. In addition, due to the fact that CL is largely researched in Western settings, it is a concern whether it can be successfully applied within Eastern settings, such as in China, where individualistic and competitive learning still dominate education at all levels (Li & Campbell, 2008; Liang, 2004; Phuong-Mai, Terlouw & Pilot, 2006). So far there has been very little research on the classroom application of CL in China.

Research Questions

The aim of the current study was to adapt CL methods for the College English teaching context in China and explore the effectiveness of specific CL techniques that suit Chinese EFL learners. Specifically studied were the effects of CL on students' English language proficiency, learning motivation and social skills, in comparison with traditional whole-class instruction. The research was particularly designed to answer the following three questions:

- 1) What is the effectiveness of cooperative learning methods in developing students' English language proficiency, compared with traditional instruction, in College English teaching in China?
- 2) What is the effectiveness of cooperative learning methods in improving students' learning motivation, compared with traditional instruction, in College English teaching in China?
- 3) What is the effectiveness of cooperative learning methods in cultivating students' social skills, compared with traditional instruction, in College English teaching in China?

CHAPTER TWO: LITERATURE REVIEW

Overview

The primary purposes of this chapter include establishing a sound theoretical framework for the application of cooperative learning (CL) in educational settings, especially in the field of second language teaching, and providing a rationale and justification for this study. This chapter consists of a review of the literature regarding the history of CL, the fundamental theoretical underpinnings of CL, major CL methods, basic elements of CL, and its positive outcomes. This chapter also elaborates on key issues in implementing CL in the classroom and the cultural appropriateness of CL in China. In accordance with the specific context for this research, this chapter carefully addresses the relevance of CL to second language (L2)/foreign language (FL) teaching, tertiary education, and large-class instruction. This is followed by a description about recent studies on using CL in L2/FL teaching around the world, as well as in China. This extensive literature review locates a gap in the existing studies—the effectiveness of CL in teaching English as a foreign language (EFL) to tertiary education learners in China— which constitutes the focus of this research.

History of Cooperative Learning

Cooperation has been an essential strategy for survival and development throughout human history. The equivalent of the old saying “Two heads are better than one” can be found in almost any language and any culture. In a similar vein, CL has a long history, dating back to the late eighteenth century, when Joseph Lancaster and Andrew Bell used and disseminated this form of instruction in England. In 1806, the concept and practices of CL were introduced to the United

States with the opening of a Lancastrian school in New York (Johnson, Johnson & Holubec, 1994; Johnson et al., 1998) and started their long journey in the United States. In the following years of the early nineteenth century, the use of CL in the classroom was highlighted to promote the educational goals at that time, such as the Americanization of a diverse student body and effectiveness in teaching a class involving mixed grades.

Colonel Francis Parker was one of the most influential proponents of CL in the late nineteenth century. His fame and success originated in his suggestions of links between CL and democratic education, his enthusiasm to advocate for the use of CL in public schools, and his efforts to spread CL perspectives and practical procedures. His methods of structuring cooperative groups influenced American education through the turn of the century. In the early twentieth century, John Dewey (1924) developed Parker's connection between CL and democracy and extended the use of CL in his project method of instruction at school. Dewey maintained that building up a democratic and cooperative setting at schools is vital for individuals to be cooperative and live democratically in real life.

Although the increasing emphasis on interpersonal competition made competitive and individual learning popular during the period from the 1930s to the early 1960s, CL advocates never lost their confidence and interest in the role of cooperation in education. During this time, Kurt Lewin (1935, 1948) and Morton Deutsch (1949) further developed their views of the group as a dynamic whole and formulated a theory of cooperation and competition. Around the late 1960s, on the basis of Deutsch's perspectives, David Johnson and his brother Roger Johnson (Johnson et al., 1994, 1998) established social interdependence theory (see the section on social interdependence theory, pp. 18-19), initiated teacher training programmes on CL, and established the Cooperative Learning Centre at the University of Minnesota. CL regained strength in the 1970s as a well-recognized effective school practice for "providing students of different ethnic groups with opportunities for nonsuperficial, cooperative interactions" (Slavin, 1995, p. 51). Since then studies on CL have abounded and some leading CL researchers (e.g. Elliot Aronson, Lynda Baloche, Elizabeth Cohen, Robyn Gillies, George Jacobs, David Johnson, Roger Johnson, Spenser Kagan, Shlomo Sharan, Yeal Sharan, Robert Slavin) have engaged in exploring specific

applications of CL to classroom teaching, which has resulted in a number of different methods and models, which will be elaborated upon later in the section on types of CL methods. In 1979, the International Association for the Study of Cooperation in Education (IASCE) was established in the United States, which aims at providing support for CL practitioners as well as disseminating advanced research on CL. Its annual conferences have served as important occasions for sharing ideas and experiences among CL educators.

Theoretical Roots of Cooperative Learning

Any effective instructional practice must have solid theoretical foundations and scientific supports. There are mainly five theoretical perspectives underlying and guiding research on CL: social interdependence theory, cognitive developmental theory, motivational theory, behavioural learning theory and humanistic psychology. These five theoretical roots provide insight into why students in cooperative groups learn more effectively and happily than those exposed to traditional teacher-fronted classroom instruction.

Social Interdependence Theory

The social interdependence theory was derived from the Gestalt School of Psychology (Lewin, 1935) and the theory of cooperation and competition (Deutsch, 1949). The Gestalt School of Psychology holds that the essence of a group is the interdependence among members that results in the group as a “dynamic whole”, which means a change in the state of any group member influences the state of the others. The theory of cooperation and competition focuses on three goal structures: cooperative, where each individual’s goal-oriented efforts contribute to others’ goal accomplishment; competitive, where each individual’s goal-oriented efforts inhibit others’ goal accomplishment; and individualistic, where individuals’ goal-oriented efforts have no consequences for others’ goal accomplishment.

Social interdependence theory, developed by David Johnson and Roger Johnson in 1970s, asserts that the “way social interdependence is structured determines how individuals interact which, in turn, determines outcomes” (Johnson et al., 1998, p. 3:6). There are three types of social interdependence relations: positive interdependence, negative interdependence and absence of interdependence (Johnson & Johnson, 1974; Johnson et al., 1998). Positive interdependence is linked to cooperation and promotive interaction where individuals encourage and facilitate each other’s efforts for success. Negative interdependence is related to competition and oppositional interaction where individuals discourage and impede each other’s efforts to achieve. The absence of interdependence is linked to individualistic efforts where individuals work independently without interaction with each other.

Since 1898, there have been over 700 studies on cooperative, competitive and individualistic learning, which provide substantial evidence that CL is superior to the other two, in that it can generate greater effort to achieve, more positive relationships among students and greater psychological health (Johnson et al., 1998). Johnson and Johnson (1994a, 1994c) also hold that too many teacher-fronted instructional practices encourage students to feel negatively interdependent with their peers and impede learning results. “Learning Together”, also known as the “Johnson Approach”, is aimed at generating effective learning through positive interdependence among learners.

Cognitive Developmental Theory

Cognitive developmental theory is mainly based on the theories of Jean Piaget (1959) and Lev Vygotsky (1978; 1986). They both emphasize the importance of peer interaction and the value of social context for generating cognitive development and effective learning. The work of Piaget is based on the premise that the cognitive disequilibrium created by social-cognitive conflicts during social interaction stimulates perspective-taking ability and cognitive development. It is argued that during cooperative efforts, participants engage in discussions where cognitive conflicts occur and are resolved, and inadequate reasoning is exposed and improved. From Piaget’s perspective,

social-arbitrary knowledge such as language, values and morality which are culture-specific, can only be acquired through social interaction with others. Translated into the context of language teaching, this perspective means it is essential to build up interactive classroom environments where students are engaged in real-life situations. Vygotsky espoused the view that knowledge is social and successful learning is constructed through cooperative efforts, where group members exchange information and insight, provide scaffolding, and help each other improve. Vygotsky's theoretical construct of the zone of proximal development (ZDP) is based on the assumption that collaborative activities with adults or more capable peers will promote cognitive growth and less competent children can benefit from peer interaction and collaborative activities.

In addition, there are controversy theories, cognitive elaborative theories and cognitive restructuring theories. Controversy theories posit that confrontation with opposing views and perspectives creates uncertainty and conceptual conflicts, which, in turn, lead to reconceptualization and cognitive development (Johnson & Johnson, 1992). Elaborative theories state that peer interaction and cooperation often produce some sort of cognitive restructuring, peer tutoring, elaborating or explaining, and the students who gain most are the more competent learners who provide these elaborated explanations and tutoring (Slavin, 1995, 2000). Similarly, cognitive restructuring theorists hold that it is only through cognitively rehearsing and restructuring learning materials that learners can retain them in memory and further integrate them into an existing cognitive structure (Wittrock, 1990).

Motivational Theory

Learning motivation theorists assert that the teacher is almost the only source of reinforcement for positive learning behaviours in the traditional teacher-fronted classroom, where students tend to feel negatively interdependent and compete against each other for reinforcement from the teacher in such forms as praise and grades. This competitive learning environment may create peer norms that oppose academic efforts' since one student's success reduces others' chances for success and thus academic efforts are not encouraged by peers. By contrast, in the CL classroom, students

form a mutual internal source of positive reinforcement for one another because of their relationship of positive interdependence (Baloche, 1998; Dishon & O’Leary, 1994; Dörnyei, 1997; Johnson et al., 1998; Kagan, 1994; Slavin, 1995, 2000).

Motivational perspectives on CL focus on three elements: goal structures, reward structures and group dynamics (Dörnyei, 1997; Johnson et al., 1998; Slavin, 1995, 2000). It is believed that “cooperative goal structures create a situation in which the only way group members can attain their own personal goals is if the group is successful” (Slavin, 1995, p. 16). Thus group members strive to help one another, and more importantly they encourage each other to make their maximum effort. The reward structure of CL is mostly linked to group reward, which means students are assessed as a group based on group performance or the sum of individual performances. Although group rewards are typically considered extrinsic motivators, “building in external reasons for students to cooperate can lead to internal motivation to work in groups” (Dishon & O’Leary, 1994, p. 58). In other words, students who are constantly extrinsically motivated to learn in cooperative groups are likely to exhibit intrinsic motivation for learning, especially with their improvement in self-esteem, peer relations, proacademic norms, and liking of class and school (Slavin, 1995, 2000). Dörnyei (1997, p. 484) states that the innovation and strength of CL primarily are derived from the “conscious and systematic exploitation of the principles of group dynamics”, which is closely linked to positive interdependence and individual accountability and conducive to strong group cohesiveness.

In addition, researchers contend that some features of CL teamwork provide a means of promoting learners’ intrinsic motivation (Jacobs & Goh, 2007). These features include students’ satisfaction from helping others and being part of group effort, their increased sense of control over and ownership of their own learning, and the use of peer evaluation and criterion-referenced assessment. There is a body of contrastive research that suggests that, in comparison with those in traditional classrooms, students in cooperative groups are seldom absent from class, they feel that their classmates want them to learn and support each other, and consequently they are motivated to try their best (Brown & Thomson, 2000; Jacobs & Goh, 2007; Slavin, 1995, 2000).

Behavioural Learning Theory

The behavioural learning theorists stress the critical role of group external reinforcers and extrinsic rewards in stimulating desirable actions. It is assumed that actions followed by extrinsic rewards are very likely to be repeated and increased, and cooperative efforts tend to be powered by extrinsic motivation to achieve group rewards. Researchers (Boud, Cohen & Sampson, 2001; Brown & Thomson, 2000; Gillies, 2007; Slavin, 1987, 1995, 2000) hold that extrinsic rewards and feedback should be given close in time to participants' performance on the basis of a well-defined behaviour standard, so that participants can be well aware of what behaviours are to be reinforced or adjusted and thus make proper improvements. A good number of studies have established the critical role of interpersonal reinforcers and punishers in affecting students' social behaviour and academic performance (Dishon & O'Leary, 1994; Johnson et al., 1998; Jolliffe, 2007).

The behavioural learning theory emphasizes the importance of group contingencies (Slavin, 1987, 2000), which means groups of students are rewarded on the basis of the behaviour of all of the group members or occasionally a single or certain members. Two elements are essential for group contingencies: group reward and individual accountability, which means "group members must be aware of the individual contributions if they are to be able to apply the interpersonal sanctions held to be central to the effectiveness of the group contingency" (Slavin, 1987, p. 34). There is a body of research to support the idea that group contingencies are particularly effective in improving student academic achievement because of the individual-performance-based group reward (Johnson et al., 1994; Slavin, 1987, 1995, 2000). This reward provides an incentive for the group members to facilitate each other's learning and help its members master learning materials so that the whole group can succeed.

Traditional CL only emphasizes cooperative interaction from a social psychological or humanistic background, with the use of group contingencies considered optional rather than essential. Some researchers even disagree with the use of contingent rewards because of the concerns that contingent rewards may undermine students' intrinsic motivation and provoke feelings of unfairness (Jacobs, 2006; Kagan & Kagan, 2009). However, Slavin (1987, p. 33)

proposed that “What is critical in cooperative learning is the combination of group contingencies and high-quality peer interaction”, which should integrate the elements of group rewards, individual accountability and cooperative interaction. Therefore, in spite of the argument that the use of extrinsic rewards should not be encouraged due to the common belief in the part of intrinsic motivation, Slavin (1987, 1995, 2000) emphasized the need for extrinsic group rewards to motivate members to learn in CL groups, especially at the initial stage or for young learners.

Humanistic Psychology

A central assumption of humanism is that human beings behave out of intentionality and values. Humanism theorists focus on human freedom, dignity, potential and autonomy, and thus give primacy to the study of human needs and interests. They value the pedagogical approach which provides a foundation for personal growth and development so that learning will continue throughout life in a self-directed manner (DeCarvalho, 1991). In the affective domain of education, humanists highlight the “uniqueness of each individual and the need for self-initiative as opposed to one-size-fits-all and teacher-fronted instruction” (Jacobs, McCafferty & DaSilva Iddings, 2006, p. 16). A primary purpose of humanism could be described as the development of self-actualized, autonomous people. In humanistic education, learning should be student-centred and personalized, and the educator’s role is that of a facilitator.

Specifically speaking, humanistic learning theories espouse the following basic principles (DeCarvalho, 1991; Gage & Berliner, 1998; Rogers & Freiberg, 1994). First, instructions should be learner-centered, based on the learners’ needs, interests and academic levels. Students are most motivated to learn what they want and need to know. Second, knowing how to learn is more important than acquiring a lot of knowledge. Primary attention should be paid to the incorporation of empowering learning experiences and the development of learners’ regulatory systems, which involve self-direction, self-reflection, self-evaluation, learning autonomy, decision making, intrinsic motivation and creativity. Third, students learn best in a non-threatening environment. In other words, creating a learning environment which is safe psychologically, emotionally as well as

physically, is an integral part of a curriculum. Fourth, facilitative teaching and group work are advocated. As a facilitator to learning, teachers provide a variety of scaffolding to cater for students' individual needs, and allow students to decide on their learning tasks or activities. Group work enables participants to develop a wide range of academic, affective and social skills, which are essential for their personal growth and future career.

The humanistic learning approach and the CL approach overlap in many areas. First, CL is centered on learners, who have the freedom to make decision concerning cooperative styles, group mates, and individual roles. Second, team tasks designed by either teachers or students themselves demonstrate different levels of difficulty, themes and purposes to suit students' needs. Third, CL practitioners also consider it critical to use self-evaluation, self-reflection as well as group processing. Fourth, it is essential for CL to create a non-threatening and supportive learning environment, with peers being positively interdependent. Obviously, all these principles incorporated in CL demonstrate its close connection to humanistic learning theories.

Types of Cooperative Learning Methods

Recent years have witnessed an increasing number of studies of CL for the purpose of improving classroom instruction; however, a considerable amount of published research “does not necessarily testify to the way CL methods are actually implemented in classrooms” (Sharan, 2002, p. 106). This calls into question the quality and adequacy of the research. This is because there are a variety of CL methods which “share the idea that students work together to learn and are responsible for their teammates' learning as well as their own” (Slavin, 1995, p. 5). Different CL methods overlap but are not equivalent in terms of their components, procedures or their particular appropriateness for subject areas, school levels and educational settings. Generally speaking, there are six major CL methods: Learning Together created by David and Roger Johnson (1994a, 1994b), Student Team Learning by Robert Slavin (1994) and his colleagues at Johns Hopkins University, Structural Approach by Spencer Kagan (1994), Jigsaw by Elliot Aronson (1997) and

his colleagues in Austin, Texas, Group Investigation (GI) by Yeal Sharan and Shlomo Sharan (1994), and Complex Instruction by Elizabeth Cohen (1994) and her colleagues.

Learning Together

Learning Together provides a generic framework for applying CL in any subject area to learners of any age, which emphasizes the integrative use of three types of CL styles: informal CL, base group CL and formal CL. Informal CL refers to having students work together in temporary, ad-hoc groups and serves as a valuable aid for students to process their learning materials effectively during direct teaching. For instance, informal groups can be used to “help set expectations as to what the lesson will cover, ensure that students cognitively process the material” being taught, and “provide closure to an instructional session” (Johnson et al., 1998, p. 1:8). Base groups are long-term, heterogeneous CL groups with stable membership, aiming at providing constant support and motivation that group members need to achieve educational success instead of working together on a specific learning tasks or assignments. Base group members are like good friends, and comprise a supportive learning community with an obligation to help each other in the academic field (Jacobs, 2006).

Formal groups, as the heart of the CL classroom, are often carefully formed according to certain principles (e.g. maximum heterogeneity), and aim at students achieving mutual learning goals through completing assigned tasks with group members cooperatively. It is assumed that any lesson or assignment may be reformulated to be cooperative. In Learning Together, teachers follow five major steps. First, they specify the objectives for the lesson either in terms of academic areas or cooperative skills. Second, they make a number of pre-instructional decisions on grouping students, assigning individual roles and tasks, and planning materials. Third, they specify the task and the positive interdependence, including explaining the learning task, structuring positive interrelations and individual responsibility among group members, then clarifying criteria for success and desired behaviour. Fourth, they monitor students’ learning and intervene within the groups, providing assistance in terms of academic knowledge and cooperative skills. Five, they

evaluate students' learning and help students assess how well their groups work together. These five steps compose a general procedure and guideline for designing and conducting the whole process of cooperative lessons.

Student Team Learning

Student Team Learning methods stress the use of group goals and group success, which cannot be achieved until all group members have grasped the materials being taught. There are three central elements in Student Team Learning methods: team rewards, individual accountability, and equal opportunities for success. Successful teams earn their team rewards when their team performances are above pre-set assessment criteria. The overall performance of each team depends on the individual performance of all teammates on the assessment (e.g. quizzes or academic games) that students take individually without help from others; in this way, each member has his/her individual accountability for team success. Team rewards and individual accountability effectively engage teammates not only in working hard to get themselves prepared for a quiz but also in helping each other to make sure every member can do well. "Equal opportunities for success means that students contribute to their teams by improving on their own past performance" (Slavin, 1995, p. 5); in other words, the more improvement points teammates gain, the more likely their team will succeed. This element allows students of different academic levels to be equally challenged to make a contribution to their teams because they compete with themselves rather than with others. Research suggests that "if students are rewarded for doing better than they have in the past, they will be more motivated to achieve than if they are rewarded for doing better than others" (Slavin, 1995, p. 5).

The two principal Student Team Learning methods which are widely used include Student Team-Achievement Divisions (STAD) and Team-Game-Tournament. The main idea behind the two methods is to motivate teammates in heterogeneous groups to help each other grasp the learning materials presented by the teacher. STAD can be used with students of all ages from elementary to tertiary level and is adaptable to most subject areas. In STAD, the cycle of

instructional procedure involves five steps. First, the teacher gives a class presentation on learning materials. Second, students work together in their teams on the materials they just learned. Group work usually “takes the form of students discussing problems together, comparing answers, and correcting any misconceptions if teammates make mistakes” (Slavin, 1994, p. 6). Third, following a couple of periods of teacher presentation and team practice, students take individual quizzes without teammates help. Fourth, individual improvement scores are calculated by comparison with their individual base scores, which indicate the level of their average academic performance in the past. Last, individual improvement scores of all teammates are added up to form the team scores and those teams whose scores exceed a certain criterion are recognized as high-performing teams with team rewards. Team-Game-Tournament, which is usually used for teaching mathematics at the elementary level, employs the same procedure as STAD apart from the replacement of quizzes with academic tournaments, in which “students play academic games with members of other teams to contribute points to their team scores” (Slavin, 1995, p. 6).

In addition there are another two Student Team Learning methods designed for particular subjects at particular grade levels: Team Accelerated Instruction for teaching mathematics at the elementary level, and Cooperative Integrated Reading and Composition for teaching reading and writing in the upper elementary level. In contrast to STAD and Team-Game-Tournament, these two methods involve individualized instruction for students of different academic levels. Specifically speaking, students are grouped in terms of their academic level and work on materials appropriate to their levels, while the teacher presents lessons which are adapted to each group’s level. Thus Team Accelerated Instruction and Cooperative Integrated Reading and Composition do not fit in with the common curricula which typically provide single-pace instruction for the class.

The Structural Approach

The Structural Approach focuses on the use of a variety of CL structures in processing learning materials. The basic premise of the Structural Approach is that “interactions in the classroom have a profound effect on the social, cognitive, and academic development of students”, therefore

“teachers should be provided with the means to direct the interaction of students in ways that will result in a range of learning outcomes” (Kagan & Kagan, 1994, p. 115). Structures are defined as distinct ways of organizing the interaction of students in the classroom. Each structure involves well-prescribed and easy-to-follow steps, which are alternatively termed as “elements” by Kagan in the Structural Approach. Kagan (1994) has provided about 100 structures aimed at six different functions: teambuilding, class-building, mastery, information exchange, communication skills, and thinking skills. Within each category of function, there are numerous structures which have different predictable outcomes in the academic, cognitive and social domains. A teacher with adequate knowledge of a variety of structures is capable of choosing the most effective structures for a desired educational goal. This can be illustrated by three structures under the category of mastery called Pairs Check, Flashcard Game and Numbered-Heads-Together (NHT).

In Pairs Checks, students work in pairs within foursomes, they take turns to solve a problem while the other coaches, and after every two problems the pair checks to see if they have the same answers as the other pair. This structure works well for acquiring new skills. In Flashcard Game, students use flashcards with questions written on one side and answers on the other, they take turns to be the tutor and the tutee to check each other’s answers. This structure is designed to facilitate the memorization of facts. In contrast, NHT, consisting of four elements (i.e. students are numbered off, they work on a given task in team, then the teacher chooses a number, and all students with that number report their teams’ work), has multiple uses such as mastering learning materials, reviewing before a test, creating an anticipatory set for a lesson and so on.

In addition, a structure can be as simple as a two-minute three-step pair work (e.g. Think-Pair-Share involving students thinking individually on a topic, then discussing in pair and last sharing their ideas with the class) or as complex as Co-op Co-op. This complex structure, designed for organizing the classroom around cooperative projects, involves ten steps: student-centred class discussion on the learning topic, selection of student learning teams, teambuilding and cooperative skill development, team-topic selection, mini-topic selection for each member, mini-topic preparation, mini-topic presentations, preparation of team presentations, team presentations, and team evaluation. This structure involves the participation of the whole

class and the structured activities may last as long as a number of weeks.

In addition to the structure, there are another three important concepts in the Structural Approach: content, activities and lessons. A structure is a format for working, or the “content-free ‘how’ of instruction” for the “social interaction of the classroom” (Kagan, 1994, p. 5:2). The content is the “what” of instruction, which fits into a structure and forms a specific and content-bound activity. Plugging different content into a structure generates different activities. A CL lesson is composed of a series of structured activities appropriate for different purposes. An experienced teacher in CL is fluent in many structures, competent in designing multi-structural lessons, and able to create their own structures by reforming individual elements according to their specific teaching discourse and objectives.

Jigsaw

Jigsaw was first designed in 1970s by Aronson and his colleagues, as an attempt to implement the desegregation of schools and build up good relations between children in multiracial situations. Its name derives from “the metaphor of putting together the pieces of a puzzle to create a whole picture” (Clarke, 1994, p. 36). The use of Jigsaw in the classroom “curbs some of the undesirable aspects of excessive competition and increases the interest children have in cooperating with one another” (Aronson & Patnoe, 1997, p. 14). The major vehicle of Jigsaw to make teammates positively interdependent is through task specialisation within the team, which makes each member and his/her work valued by the others. Jigsaw fits best in the situations where learning is based on text-based materials “that can be divided equally among students” (Aronson & Patnoe, 1997, p. 25) and each particular section of the text is distributed to only one particular member of the home team.

Jigsaw typically involves three steps: first, students are divided into heterogeneous home groups, with each member assigned a particular section of the learning unit to study; second, students focusing on the same sections meet together in focus teams to explore the particular aspects; third, students return to their home groups to share with each other what they have learned

in their focus groups so that everyone gets a whole picture of the learning unit. Slavin (1995) creates Jigsaw II by introducing two elements to Aronson's Jigsaw (which was renamed Jigsaw I by Slavin). The Original Jigsaw (or Jigsaw I) requires that learning materials be divided neatly into sections that are comprehensible by themselves. However, this rarely happens since all sections in a text are usually intertwined. Thus teachers usually need to restructure learning materials, which greatly adds to their workload and thus make Jigsaw I less practical. In order to make use of existing textbooks, Jigsaw II proposes that all students read the whole text but each team member be assigned a particular aspect to work on in the focus team. Second, home team recognition is introduced in Jigsaw II based on the average performance of all team members on a quiz which takes place after teamwork, as in the other Student Team Learning methods.

However, there are some concerns that elements of CL might be missing in Jigsaw methods. A common issue is how to make students positively interdependent and sufficiently enthusiastic to help each other's learning in focus groups. This issue may become more critical in Jigsaw II where home teams' success is recognized. Home team recognition effectively consolidates the home team cohesion; but on the other hand, it is likely to build a competitive spirit between home teams and inhibit the mutual support within focus teams.

Group Investigation

Group Investigation (GI) has its origins in Dewey's philosophy of education, which views cooperation in the classroom as a prerequisite for dealing with the complex issues of the real world and emphasizes the importance of investigation in learning since "meaningful learning proceeds through the steps of scientific inquiry, whereby students experience how knowledge is generated" (Sharan & Sharan, 1994, p. 98). In contrast to the Jigsaw, the Structural Approach and Student Team Learning methods "which are oriented toward student acquisition of predetermined facts and skills" (Kagan, 1994, p. 19:10), GI is designed for more diverse and complex learning experiences which allow students a greater amount of freedom and autonomy in the learning process. The effective implementation of GI requires, however, that students have adequate ability

to plan and study together and basic skills in team cooperation and communication; so it is suggested that GI should not be introduced until “students are used to working together to achieve academic goals” (Sharan & Sharan, 1994, p. 97).

In GI, students in their teams work on a broad multifaceted problem through cooperation with teammates and other teams. It involves six consecutive stages: First, students determine the general topic and subtopics for investigation, and are organized into research teams. Second, teams plan their investigations, involving dividing the work, assigning roles and specific research questions, as well as seeking resources. Third, teams carry out their investigations through multilateral interaction and interpretation with teammates, other teams, the teacher and other resource persons. Fourth, teams plan their presentations as the final report of their group investigations. Fifth, teams present their final reports with each teammate taking an active role in the presentation. Last, the teacher and students evaluate their projects.

There are four basic features of GI: investigation, interaction, interpretation, and intrinsic motivation. First, investigation is the general means for learning in GI; in other words, an inquiry community of students actively constructs knowledge through investigating a challenging multifaceted problem and searching for potential solutions. Second, effective interaction is called for all through the process of GI. By intellectual and social interaction, students share information, encourage each other, provide scaffoldings for each other’s work, settle disagreements, and rework their personal knowledge based on the new acquisition. Third, when students interact with each other, they are actually in the process of interpreting their combined findings and knowledge through discussing, clarifying, supplementing and synthesizing ideas. This is conducive to deepening their understanding, generating new knowledge and enhancing learning outcomes. Fourth, students are intrinsically motivated to learn because they play active roles throughout their learning: they decide what and how to learn according to their current interest and needs, they set guidelines to act upon, and thus have a great deal of control over their learning. In addition, intrinsic motivation is more likely to be generated through the process of exploring the interesting multi-faceted problems which demand more complex enquiry and have no clear predictable answers (Brown & Thomson, 2000).

Complex Instruction

Complex instruction (Cohen et al., 1994) is aimed at building respect for all the intellectual abilities students have, and is appropriate in linguistically and academically diverse learning settings, particularly in “bilingual education and heterogeneous classes containing language minority students” (Slavin, 1995, p. 11). Its major focus is on addressing the issue of students’ unequal influence on and participation in the group task due to status problems, that is, group work is usually dominated by high-status students while low-status students, who are expected to be inferior linguistically, academically or socially, are likely to be ignored. This situation can eventually result in serious learning problems.

One critical solution to status problem is through the multiple-abilities treatment, which is “based on the teacher’s public recognition of a wealth of intellectual abilities that are relevant and valued in the classroom and in daily life” (Cohen et al., 1994, p. 85). The teacher should make it clear to students that each of them has his/her own strengths and weaknesses among the different valuable abilities. However, successful multiple-abilities treatment lies in the design and implementation of multi-abilities group tasks which call for different intellectual abilities with special consideration of those relevant to low-status students. In other words, it is essential for the teacher to assign competence to low-status students through the process of designing, implementing and assessing group work by recognizing that certain ability demonstrated by a low-status student is very important to complete the group task. This technique is aimed at raising not only low-status students’ self-concept but also others’ expectation for them, which consequently leads them to greater access to active interaction and learning.

In addition, Cohen and her colleagues propose two necessary supporting elements for Complex Instruction. One is building up a special classroom management system, which includes preparing activity cards with task instructions, requiring individual reports, assigning individual roles and developing cooperative norms. The other is establishing specialized curricula which allow adaptation of existing materials to Complex Instruction and particularly entails the design of multiple-abilities open-ended learning tasks.

Basic Principles and Elements of Cooperative Learning

Although the six major CL methods entail different components, procedures and educational merits, there is one fundamental goal they all share, that is, building a positively interdependent learning environment where students work in groups with dual responsibility for their own and each other's learning. This common goal results in some basic principles and elements that all CL methods value and emphasize. Many CL researchers argue that it is these basic principles and elements that define CL as distinct from other types of group work. In other words, the absence of any required elements or principles may cause CL to be less effective. A review of the perspectives on the basic principles and elements posited by different researchers (e.g. Aronson & Patnoe, 1997; Brown & Thomson, 2000; Cohen et al., 1994; Dishon & O'Leary, 1998; Johnson et al., 1994, 1998; Kagan, 1994; Sharan & Sharan, 1994; Slavin, 1995) reveals that positive interdependence and individual accountability are widely accepted as the two fundamental constructs of all CL methods. Without these two principles, CL loses its identity and becomes indistinguishable from other types of group work. In addition, there are also some other principles which are considered indispensable to effective learning by many CL advocates. They include promotive simultaneous interaction, equal participation, equal opportunities for success, social skills, and group processing.

“Positive interdependence is linking students together so one cannot succeed unless all group members succeed” (Johnson et al., 1998, p. 4:7), and it enables students to reach a goal beyond individual ability and maximize their learning through the dual responsibility for both oneself and the other team members. Positive interdependence can be structured by carefully arranging mutual goals, group rewards, divisible learning resources, multiple-abilities tasks and individual roles in cooperative groups. It is essential that group members perceive that they are linked together in a way that they “sink or swim together” since no one can succeed unless everyone succeeds.

Individual accountability requires that every teammate is accountable for completing a particular task and no one can “hitchhike” on the work of others. It is important to ensure that

students know that their contribution to teamwork can be individually identified and assessed. Techniques to ensure individual accountability typically include testing students individually after group work, assigning each member an individual role or task, and randomly selecting certain students as team representative to present team work.

Promotive simultaneous interaction is a synthesis of promotive interaction (Gillies, 2007; Johnson et al., 1994, 1998; Sharan & Sharan, 1994) and the simultaneity principle (Kagan, 1994). Promotive interaction refers to students' effort to facilitate each other's success and is conducive to caring and committed relationships, psychological adjustment, social competence and low levels of anxiety and stress. The simultaneity principle means more interaction can be generated among peers simultaneously within smaller groups than larger ones. Thus, this principle is intended to engage more students in positive interaction and language production at the same time in a supportive, caring and non-threatening learning environment.

Equal participation of group members is another distinctive characteristic of CL (Aronson & Patnoe, 1997; Baloche, 1998; Cohen et al., 1994; Kagan, 1994; Sharan & Sharan, 1994), and is considered a natural result of positive interdependence and individual accountability.

An equal opportunity for success is an element particularly highlighted in Student Team Learning methods (Slavin, 1994, 1995). It can be realized through a number of techniques: grouping students to ensure within-team heterogeneity and between-team homogeneity, using improvement points, integrating competition with equals, and adapting tasks to individual performance levels. Equal opportunity for success especially benefits the low achievers, who often get negative feedback on their academic performance and feel that academic success is always beyond them in the traditional competitive classroom.

Social skills are also termed as interpersonal skills, small-group skills, cooperative skills or team skills in different CL literature (Aronson & Patnoe, 1997; Baloche, 1998; Brown & Thomson, 2000; Cohen et al., 1994; Dishon and O'Leary, 1998; Gillies, 2007; Johnson et al., 1998; Kagan, 1994; Sharan & Sharan, 1994; Slavin, 1995). It is believed that the appropriate grasp and use of social skills is essential to complete group tasks and gain academic achievements. There is evidence that students in cooperative groups who are taught specific skills achieve better in school

than do those who are not (Slavin, 2000), so teaching relevant social skills is considered essential especially for children (Brown & Thomson, 2000; Dishon & O’Leary, 1998; Gillies & Ashman, 2003a, Johnson et al., 1994, 1998; Kagan, 1994).

Group processing involves students reflecting on their learning experience and discussing how well the group work is going and what actions should be maintained or changed in order to improve the effectiveness and efficiency of the cooperative groups.

Positive Outcomes of Cooperative Learning

Brown and Thomson (2000, p. 14) contend that the “two main purposes of CL are to develop academic and interpersonal skills” so that young people can become competent individuals, face the world with confidence and know how to work well together with others in the future. Kagan (1994) suggests CL may bring about a number of positive outcomes: academic achievement especially for minority and low achievers; improved ethnic relations in integrated classrooms; and improved social and affective development, including students’ social skills, self-esteem and liking for others. According to Johnson et al. (1994, 1998), positive outcomes relating to CL can be categorized into three areas: effort to achieve, positive interpersonal relationships and psychological health. A comprehensive literature review made by Slavin (1995, p. 50) leads to the same perspective that “cooperative learning is not only an instructional technique for increasing student achievement, it is also a way of creating a happy, pro-social environment in the classroom, one that has important benefits for a wide array of affective and interpersonal outcomes.” A synthesis of the benefits of CL based on the findings of leading researchers (e.g. Brown & Thomson, 2000; Dörnyei, 1997; Gillies & Ashman, 2003a; Johnson & Johnson, 2003; Johnson et al., 1998; Kagan, 1994; McCafferty et al., 2006; Sharan, 2003; Sharan & Shachar, 1988; Slavin, 1995, 2000) comes to the conclusion that the positive outcomes of CL can be categorized into three major areas: academic achievement, social skills and learning motivation.

The conclusion that CL promotes higher academic achievement can be supported by the

comprehensive meta-analyses conducted by Johnson, Maruyama, Johnson, Nelson and Skon (1981), Slavin (1995) and Hattie (2009). Johnson et al. (1981) review 122 achievement-related studies and compare their relative effectiveness in promoting achievement. The meta-analysis points to the obvious superiority of group cooperation over interpersonal competition and individualistic goal structures. Slavin (1995) examines 99 experimental-control comparisons, which meet the inclusion criteria of two initially equivalent groups studying the same material and using the same achievement measures, with experiment duration of more than four weeks. Slavin's review reports that 64% significantly favour CL while only 5% favour the traditional approaches. According to Hattie (2009), a meta-analysis of 1,898 studies of school-age children indicated a universal agreement that CL was more effective than competitive or individualistic learning in improving learners' academic proficiency.

Social skills in the field of CL are defined as "those specific behaviours performed by all group members that help the group complete the task and appreciate each other when the task is finished" (Dishon & O'Leary, 1998, p. 93). According to CL researchers, there are a number of ways to classify social skills (for reviews, see Baloché, 1998; Brown & Thomson, 2000; Hill & Eckert, 1995; Johnson et al., 1994, 1998; Kagan, 1994; Slavin, 1995). But if we probe the elaboration and breakdowns of the social skills noted by each researcher, it turns out that they overlap each other and the majority of the components are quite similar in spite of different wording. The major social skills can be categorized into two general types: task social skills and interrelationship social skills (or alternatively termed as maintenance social skills) (Brown & Thomson, 2000; Dishon & O'Leary, 1994). Task skills refer to those content-focus behaviours to complete the task such as contributing ideas, seeking information, staying on task, checking for understanding, elaborating ideas, and following directions. Interrelationship (or maintenance) skills refer to those behaviours to sustain and develop positive peer relations, and emphasize group cohesiveness and stability to facilitate the effectiveness and efficiency of teamwork (e.g. being encouraging, showing empathy, checking for agreement, managing conflict, etc.). According to Johnson et al. (1998), there have been more than 180 contrastive studies related to interpersonal attraction and 106 studies related to social support since 1940s. These studies have presented a

clear picture that, compared with competition and individualistic effort, cooperation promotes positive interpersonal relationships and generates peer liking and social support.

Johnson and Johnson (2003) define motivation as the driving force to move people towards their desired outcomes and commit effort to achieve goals that they perceive as being meaningful and worthwhile. In cooperative groups, due to positive interdependence in terms of goals, rewards, resources, roles and tasks, learners are likely to develop a strong sense of group cohesion and to be motivated not only to move towards the goal themselves but also to encourage and help others to achieve (Dörnyei, 1997; Jacobs & Goh, 2007; Johnson & Johnson, 2003; Kagan & Kagan, 2009). The researchers assume that CL primarily promotes intrinsic motivation in the long term because learners find it fun, interesting and enjoyable to work in a supportive and safe learning environment, and therefore exert greater internal pressure to achieve. Johnson and Johnson (2003, p. 164) state that “The more co-operative individuals’ attitudes, the more they see themselves as being intrinsically motivated”, which can further increase their perseverance in pursuit of goals, joint efficacy, desire for success and joy of learning.

Key Issues in Implementing Cooperative Learning

A successful use of CL in the classroom substantively relies on how well the key issues in implementing CL are understood and addressed. These issues include five aspects: how to group students, how to make groups function as cooperative teams, how to select CL methods and techniques, how to assess CL group work, and what roles teachers should play in the CL classroom.

Grouping Students

Using CL in ELT entails appropriately grouping students with differing levels of language proficiency, in a supportive environment where all group members benefit from the interactive

experience. In other words, teams are the base and core of most CL activities. When forming groups, three factors must be taken into consideration: size, duration, and selection (Jacobs, 2006; Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994).

Group Size

At the initial stage when students are new to CL, pair work is ideal since it encourages greater participation and is easier to coordinate and manage (Jacobs, 2006; Jolliffe, 2007; Kagan, 1994). In the EFL context, pair work may generate the greatest amount of language output because students are less likely to be left out or hide away from sharing information with their partners. When both the teacher and students have gained positive experience and become comfortable with working in a CL environment, larger groups can be used, which has advantages for processing more complex learning activities and developing a wider range of cooperative skills. "Larger groups also offer the possibility of differing opinions and perspectives in relation to experience. Additionally, larger groups make it easier for teachers to monitor each of the groups in a classroom, there being fewer of them" (Jacobs, 2006, p. 32).

When choosing a larger group size, a foursome is mostly preferred and commonly recommended by teachers and researchers. This is mainly because a foursome involves within-team flexibility of being split into two pairs, being switched (forming a new pair with a different teammate), and being squared (combining two pairs together). This flexibility lends itself to the use of a good number of CL structures such as Think-Pair-Share, Listen-Pair-Square and Write-Pair-Switch-Square (Jacobs & Goh, 2007; Kagan, 1994). It is generally recommended that an effective learning group should consist of no more than four members (Jacobs, 2006; Jacobs & Goh, 2007; Jacobs & Ratmanida, 1996; Johnson & Johnson, 1994a; Kagan, 1994); otherwise it is very likely that "many of the benefits in terms of increased involvement are lost" (Jacobs & Ratmanida, 1996, p. 113). It is assumed that functioning problems come in the classroom when groups get to six or more and so they are seldom used (Brown & Thomson, 2000; Jacobs, 2006; Jolliffe, 2007, Kagan, 1994; Slavin, 1995).

Group Duration

A number of factors need to be taken into consideration when determining how long groups should stay together. First, it depends on the type of group being used. As mentioned in the section on types of CL methods, there are three types of CL group: informal CL group, formal CL group and base CL group (Johnson & Johnson, 1994a; 1994b). Informal CL groups are temporary groups, lasting as briefly as a few minutes or up to a class session. The duration of the formal group may vary from one class session to six weeks or even over ten weeks (Jacobs & Goh, 2007). Many researchers (e.g. Brown & Thomson, 2000; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007) have recommended that formal groups should not be changed too frequently in order that students have time for establishing group cohesion and sorting out problems they may encounter in working with some group members. Meanwhile, they also suggest groups should not stay intact for too long so that students have the opportunity to work with and learn from different peers, which is not only conducive to the development of their academic achievement but also their social skills. Base groups are long-term group which last for at least a year and preferably for a few years until all members have graduated (Johnson et al., 1998).

Second, group duration also depends on what CL techniques or group activities are used. As mentioned in the section on CL methods, different CL methods involve different techniques, which can provide activities lasting either as short as a few minutes or as long as several weeks. For instance, Listen-Pair-Share (Kagan, 1994) is useful for a brief task, in which students listen to a short recording first, then each member of a pair speaks for a given time—say, half a minute—and finally students share their partner's ideas with the class. Some methods or techniques, for instance, GI (Sharan & Sharan, 1994) and Co-op Co-op (Kagan, 1994), can be used to organize long and complex tasks, where students decide on the project topic to be investigated, and then each group selects a research subtopic and works together on a presentation to the whole class. Some other techniques are very flexible in structuring both short and long learning tasks. For example, NHT can be based on a three-minute group discussion during the class and then students with a certain number from each group will be selected to report and

explain their group work. In this case, the whole process may last only five minutes. NHT can be also based on an assignment to be completed within a few weeks; for instance, students work in groups to prepare a speech based on their research and investigation on a topic, and several weeks are given to students to do the research, investigation and speech preparation before certain members are selected to present to the whole class.

Third, group duration depends on the extent to which student are familiar with CL teamwork and able to work together cooperatively (Jacobs, 2006; Kagan, 1994). If students fail to cooperate adequately, teachers should focus on the team-building process for a supportive and cohesive team, rather than trying to fix things through changing groups around. It should be made clear to students that complaints about their present teammates cannot form a reason for changing groups, and they can change groups only when they have demonstrated good performance in their current groups (Brown & Thomson, 2000). For those who are new to CL, long-term groups give students “more reason to overcome difficulties they may have in working with certain group mates if they know their group will exist for weeks or months” (Jacobs, 2006, p. 35). On the other hand, when students work well in cooperative groups and exhibit enough relevant teamwork skills, groups can be changed more often so that students have more opportunities to work with a variety of peers, which is conducive to the development of both academic proficiency and social skills.

Generally speaking, there would never be fixed rules on the group duration, because most deciding factors lie in the specific teaching situations. Teachers need to be flexible in making decisions to suit their current learning materials and teaching objectives, taking into account the type of groups, techniques and learning tasks being used. The right decision at the right time for the right group of students also depends on teachers’ experience in using CL groups and considerable insight into their teaching contexts.

Group Selection and Composition

There are three major ways of selecting students to groups: random selection, student selection and teacher selection (Brown & Thomson, 2000; Jacobs, 2006; Johnson et al, 1998). Random

selection may create a group of students who do not have the necessary skills for learning tasks and are unlikely to complete tasks. As a result, it is safe to use random selection only when the task involves a low level of academic challenge. With student selection, homogeneous groups may be created since birds of a feather flock together, and this is not conducive to the development of academic competence and a wider range of social skills (especially social acceptance and tolerance). “Student-selected groups often have powerful social agendas that take up their time and attention and results in much “off-task” behaviour” (Brown & Thomson, 2000, p. 64). Teacher selection is the most popular and commonly recommended method for assigning students to CL groups, because, with many factors (e.g. teaching objectives and content, students’ academic level, gender, social class and personality) taken into consideration, teacher-selected groups are likely to achieve a maximum level of between-group homogeneity and within-group heterogeneity (Johnson et al., 1998; Kagan, 1994; Slavin, 1995). Between-group homogeneity allows for fair competition between groups while within-group heterogeneity provides more possibilities for constructive interaction among group members and enables them to benefit from diverse perspectives and skills; these are considered the two essential features in the process of forming groups. Most leading CL researchers (e.g. Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994; Slavin, 1995) believe that students generally learn best in cooperative heterogeneous groups. As Freeman and Freeman (1994, p. 154) put it, “When students work collaboratively, diversity is an asset to be celebrated since the varied experiences, knowledge, and interests students in each group bring to the task at hand add to the potential for learning.”

However, there remain some concerns about the effectiveness of assigning students of mixed academic competence into a group (Jacobs, 2006; Jacobs & Goh, 2007; Slavin, 1995), because high achievers may feel bored while low achievers are intimidated when working together. Actually, many researchers (e.g. Johnson et al., 1998; Kagan, 1994; Slavin, 1995) have found both high achievers and low achievers are likely to benefit from working cooperatively towards a mutual goal., According to cognitive elaboration theories, high achiever are actually helping themselves when helping others, because the rehearsal, elaboration and reflection involved in teaching others may also help themselves deepen understanding, polish perspectives, and more

importantly enhance their teaching and communicative skills which are very likely necessary for their future career. “There is evidence to show that high achievers do at least as well and often better in cooperative classrooms” (Brown & Thomson, 2000, p. 65). On the other hand, cognitive developmental theories, especially Vygotsky’s theoretical construct of the ZDP, emphasize the important role of peer tutoring and peer scaffolding involved in the collaborative activities, from which less competent learners can benefit a lot to promote their cognitive level and academic proficiency. In spite of this, we cannot deny the possibility that, in the context that high achievers are constantly on the giving end while the low achievers are on the receiving end, learning motivation and group effectiveness might be hampered. Thus attention should be paid to involving a wide range of intelligences (Cohen, 1994; Jacobs, 2006) and skills when designing tasks, so that every group member, either a high or low achiever, can give and receive help in some sense. In other words, high achievers not only play a role as an importance resource but also are able to seek and find challenge. Likewise, low achievers not only feel they are helped and supported by more competent peers but are also able to make their own contribution valued by peers.

A few other researchers have concerns about the effectiveness of assigning students of mixed gender into a group (Belenky, Clinchy, Goldberger & Tarule, 1997; Phuong-Mai, Terlouw & Pilot, 2006). They reported that CL may provide a less equitable and effective learning environment for female learners especially in the Confucian Heritage Culture (CHC), where there is a traditional assumption that women should be passive, reticent and inclined to follow instead of leading. However, some other researchers have reported different findings relating to CHC pedagogical contexts. Wee and Jacobs’s (2006) research with Asian students suggested that mixed-gender groups are more likely to generate different perspectives, enable students to stay on task and be committed to helping each other. Chin, Teh & Fong (1988) found the CL method was more effective for female Singaporean learners in learning mathematics although not for English language, while it was just the opposite for male learners. This suggests that gender may not impede group effectiveness because both males and females perform better than the other when using CL activities in some areas. In addition, with changes in society and the popularity of equal educational opportunity, the gap between men and women in social status is considerably smaller

in most nations, including those strongly influenced by the CHC. In China, since the founding of the Republics in 1949, men and women are entitled to equal rights in terms of social, educational, and political matters (Brick, 1991). Another interesting piece of evidence that, in China, girls are no less active in classroom communication than boys is a picture going with an article on education matters in China (Powell, 2009, p. 29). The picture, focusing on routine classroom teaching at an elementary class, vividly shows that, while boys still hesitate or feel unready to respond to the teacher's questions, several girls have already put their hands up high. Generally speaking, mixed-gender groups may not have a negative effect on CL teamwork in many CHC contexts including China.

Functioning as a Cooperative Learning Group

Once groups have been formed, the next factor to consider is what strategies can be employed to make groups work well together and enhance group functioning. Effective group functioning mainly lies in good relations (or positive interdependence) of group members and adequate cooperative skills. So it is advocated that efforts be engaged in fostering the group cohesiveness and developing students' cooperative skills.

First, most CL educators (e.g. Brown & Thomson, 2000; Gillies, 2007; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; Slavin, 1995) caution that special attention should be paid to design teambuilding activities (termed as ice breakers) at the initial stage of CL, aimed at getting group members acquainted with each other, building up team identity, and creating feelings of trust and togetherness among group members. In other words, teambuilding activities lead to a supportive and relaxed atmosphere where students feel comfortable learning together. Kagan (1994) and Jacobs and Goh (2007) have proposed numerous enjoyable and useful structures for ice breakers, such as Team Interview on each other's name (relating to how they get their name, how they like their name, etc.) and Team Project for creating their team name. It is generally assumed that social and personal interaction is the major goal of teambuilding activities, therefore their "academic challenge should not be too high so that students can concentrate on

getting along with each other” (Brown & Thomson, 2000, p. 69).

Second, some class time should be spent directly in teaching students cooperative skills which are indispensable to team success (Dishon & O’Leary, 1994; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; Slavin, 1995). This is believed to be a worthwhile time investment since it is followed by a “pay-off in smoother running of the classroom and more effective learning strategies for students” (Brown & Thomson, 2000, p. 75). However, since teaching cooperative skills is invariably time-consuming, some educators doubt whether it is wise to spend so much precious class time teaching cooperative skills (Jacobs, 2006). They feel that students may develop their cooperative skills through working together on well-designed activities. Actually, the best way of teaching cooperative skills is to integrate them into specific carefully-planned tasks and teach them in authentic learning contexts through hands-on practice under teachers’ guidance (Gillies, 2007; Jacobs, 2006). A good design of a CL task makes students feel they are obliged to work together cooperatively in order to reach the intended learning objectives. Also, students may easily become positive about using cooperative skills when they benefit from them in practice. An effective way to enhance students’ cooperative skills is assigning them individual roles (e.g. encourager, understanding checker, taskmaster, etc.) for helping the group achieve its goals, which is often accompanied by the provision of useful phrases and conversational gambits for specific roles so that they know what to say to get their cooperative intention across (Brown & Thomson, 2000; Dishon & O’Leary, 1994; Jacobs, 2006; Jolliffe, 2007; Kagan, 1994).

Another important strategy that enhances the effectiveness of group functioning is group reflection, which is one of the essential principles of CL (Dishon & O’Leary, 1994; Gillies, 2007; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; Slavin, 1995). Group reflection fundamentally involves three components: evaluating how well the group functioned (i.e. what went well or badly), analyzing why that happened and what they could have done to make it better, and finally setting new goals for a higher level of group functioning in the future. It is obvious that frequent engagement of students in this reflective process substantially improves students’ sense of collaboration and overall group functioning. However, one of the major

problems in group reflection is that students may simply provide general and vague responses (Brown & Thomson, 2000, Holt, 1993; Johnson et al., 1998; Kagan, 1994), which involve no specific useful information of what actually happened and what can be done to make improvements. For instance, when asked “What do you think hindered your group work?” students may just simply answer “We didn’t work together well”. To counteract vague responses, the teacher can provide specific evaluation criteria for students to refer to during group reflection (Abram et al., 2002; Jacobs & Goh, 2007; Joritz-Nakagawa, 2006).

Selecting Cooperative Learning Methods and Techniques

As elaborated in the section on types of CL methods, the six major CL methods include numerous techniques and structures. Selecting the appropriate methods and techniques or structures for a particular teaching context is always a critical issue for effective use of CL. Synthesizing the points of view on this issue from a variety of CL literature generates four general criteria the teacher should follow when making selections.

First of all, different methods and techniques may have different anticipated outcomes and expected educational objectives, so the teacher primarily makes selections according to their specific teaching value and aim (Dishon & O’Leary; Jacobs & Goh, 2007; Kagan, 1994; Sharan, 2002). Some CL methods or techniques are oriented towards mastery of basic skills or memorization of basic facts, while others are targeted at completion of complex team projects or higher order thinking skills. The Structural Approach provides about 100 structures of diverse functions and teaching objectives, for instance, from learning vocabulary, sharing information, and developing social skills to promote high-level thinking (Kagan, 1994).

Second, the selection is based on the length of time allocated to CL activities. Different methods and techniques may involve particular procedures of different lengths of time. For instance, the Structural Approach involves some structures (e.g. Think-Pair-Share, Flashcard Game) which require a few minutes to complete a particular activity while some methods and structures (e.g. GI and Co-op Co-op) usually need much longer—several sessions or even some

weeks—to carry out a team task. In addition, some others (e.g. Jigsaw and NHT) are quite flexible and may fit in varying lengths of time.

Third, the selection should be made according to students' age and social skills. Different methods and techniques may make different demands on the students' social skills. Students who are very young or weak in social skills should be exposed to highly-structured techniques or methods (e.g. Student Team Learning, and many structures in the Structural Approach), which specialize in organizing team tasks involving well-designed learning materials with clearly-defined procedures as well as the integration of extrinsic rewards (Brown & Brown, 2000; Jolliffe, 2007; Kagan, 1994; Sharan, 2002; Slavin, 1995). On the other hand, the teacher can select methods and structures involving complex project designs (e.g. GI and Co-op and Co-op) for students who have better social skills and function well in group work.

Fourth, teachers' familiarity with CL methods and techniques and their expertise in using them should also be taken into consideration when making selections. Kagan (1994) suggests that teachers should start from some simple structures included in the Structural Approach, like Think-Pair-Share and Roundtable, which involve relatively rigid ways of structuring the classroom and can fit into any stage of a lesson design. It is recommended that teachers new to CL make detailed lesson plans and use short activities based on simple structures (Jolliffe, 2007; Joritz-Nakagawa, 2006). Repeated practice of the same structure will “smooth out the rough edges” both on the part of teachers and students (Jacobs & Goh, 2007, p. 31). As teachers become comfortable with simple structures and feel competent in the art of managing a classroom of teams, they move on to complex techniques and methods involving more procedures and a longer process. There is always some trial-and-error experimentation with CL before teachers gain adequate expertise in selecting appropriate CL models that best fits their own style.

Fifth, the selection also depends on the existing curricular and subject areas. A number of CL methods are particularly designed for certain curriculum content or subject areas, so these methods can only be used in a limited way when certain requirements are met. For instance, Team Accelerated Instruction requires a curriculum which allows for individualized instruction for students of different academic levels within a class, and also this method is specially intended for

maths teaching at the elementary level (Slavin, 1995). Jigsaw is particularly suitable for learning which is based on the text-based materials (Aronson & Patnoe, 1997). The use of Complex Instruction works with team tasks which are open-ended involving multiple abilities, and particularly suits dual-language settings (Cohen et al., 1994).

It is also very important to note that a CL lesson is often a combination of different CL methods which serves for varying teaching objectives (Holt, 1993; Kagan, 1994; Sharan, 2002). More often than not, a ready-to-use CL method or structure is not available for a particular learning setting, so teachers should know how to make adaptations and modifications on the existing CL methods or structures to fit in a specific teaching context. “Cooperative methods grow out of the modifications and adaptations made by professional educators in response to the unique demands of their teaching” (Holt, 1993, p. 3).

Assessing Cooperative Learning Group Work

Assessing group work is an integral part of the CL process because students reflecting on their performance in teamwork is universally considered to be one of essential elements of CL. Johnson et al. (1998, p. 8:6) state that in CL groups, “students learn almost as much from assessing the quality of their own and their classmates’ work as they do from participating in the instructional activities”.

Generally speaking, assessment in education can be divided into two types: summative and formative (Boud et al., 2001; Harmer, 2007; Johnson et al., 1998), or sometimes alternatively termed static and dynamic (Falsgraf, 2009). Static summative assessment is rooted in the positivist assumption that a relatively stable knowledge state exists and can be measured through testing techniques which elicit and analyze evidence of that knowledge. It often takes the form of one-off measurements, ranging from a large-scale public standardized examination to a term quiz (Harmer, 2007). A common feature shared by all forms of summative assessment is that they solely focus on assessing learning outcomes and providing specific grades or scores as an indication of learners’ current levels of achievement or proficiency. On the other hand, dynamic formative assessment

derives from the interpretivist assumption that learning is complex and individualized and cannot be judged by a one-off measure. So formative assessment focuses on assessing learning process, providing interpretation, feedback and comments from both teachers and peers as a course is progressing, aimed at helping learners know their present state of learning and how to improve their learning performance (Falsgraf, 2009). Falsgraf has further advocated that each type of assessment has its own limitations and thus achieving a balance between them can not only improve power and accuracy in measuring students' learning outcomes but also enhance their learning sense and learning performance.

Assessment strategies used with cooperative groups are mostly a combination of formative assessment and summative assessment, with the former as the foundation of the latter (Abram et al., 2002; Gillies, 2007; Jacobs, 2007; Johnson et al., 1998; Kagan, 1994; McCafferty et al., 2006). In other words, the final products of group work are often graded or scored based on the relevant feedback and comments from teachers and peers. Assessment procedures in CL often involve the following general steps: students are assigned in groups, working out a group product (e.g. presentation or composition on a topic), or preparing for a test together; and then students' performances are assessed either as a group or individually, which involves not only giving specific grades or scores but also integrating immediate clarification of weaknesses and further providing immediate suggestions for remediation (Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; Slavin, 1995).

Assessing students in terms of group outcomes or giving group grades is a very important strategy to maintain group members' positive interdependence because group members sink or swim together (Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Joritz-Nakagawa, 2006; Kagan, 1994; Slavin, 1995). There are two main ways of grading: norm-referenced and criterion-referenced. By norm-referenced grading, the score of one student may affect the grades of others (Boud et al., 200; Jacobs & Goh, 2007). For instance, if one student receives a score of 75 and the average score is 85, this student's grade could go down to a C. With norm-referenced grading, students are measured against each other and half of them are destined for grades below average in theory (Bracey, 2006). Thus this grading system "may foster competition among

learners, because if students help others learn more and score higher, these helpful students could be lowering their own grades” (Jacobs & Goh, 2007, p. 36). In contrast, criterion-referenced grading means that one student’s score has no impact on the grades of others, because this grading system “would measure people along a continuum of achievement against specific criteria” (Bracey, 2006, p. 128). In other words, students are graded entirely according to their own performances against a list of criteria and they compete with themselves rather than others. In this way, students do not feel apprehensive about helping others, and this greatly facilitates cooperation and promotive interaction in group work. Therefore, there is a universal agreement that criterion-referenced grading system is employed when assessing groups’ performances and achievements in CL (Boud et al., 2001; Holt, 1993; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; McCafferty et al., 2006; Slavin, 1995). Moreover, empirical studies have shown evidence that the presence of clear and accessible grading criteria improves the quality of group products (Joritz-Nakagawa, 2006), and also enables CL groups to spend significantly more time on-task, discussing learning content and evaluating group products; in consequence, this substantially improves the academic nature of group discussion, the quality of feedback from teachers as well as group and individual learning outcomes (Abram et al., 2002).

Although most educationists in the field of CL support the use of group grades, there still exist some concerns that group grades may not provide a reliable and fair measure of students’ work (Jacobs & Goh, 2007; Johnson et al., 1998; Slavin, 1995). For instance, if two students of equal proficiency are assigned to groups of different levels, one having more capable group mates than the other, it is very likely that the student in the more capable group receives a higher grade. Researchers and experts have proposed several solutions to this problem. First, CL groups should be formed on the basis of not only within-group heterogeneity but also a maximum of between-group homogeneity, which means that groups should be of a similar academic level at the starting point (Dishon & O’Leary, 1998; Holt, 1993; Jacobs & Goh, 2007; Johnson et al., 1998; Jolliffe, 2007; Kagan, 1994; McCafferty et al., 2006; Slavin, 1995). Second, researchers suggest using non-grade rewards such as certificates or other types of recognition for excellent group work (Jacobs & Goh, 2007). It is assumed that group grades or rewards can be dropped when students

find involvement in CL to be intrinsically satisfying and they work together well without group grades or rewards as external motivators (Boud et al., 2001; Wee & Jacobs, 2006). Third, Slavin (1995) proposed that group grades be calculated by averaging improvement points gained by individual group members. Improvement points refer to the sum of scores by which a student improves over his/her initial base scores indicating the starting-point academic level. The purpose of improvement points is to “make it possible for all students to bring maximum points to their teams, whatever their level of past performance” (Slavin, 1995, p. 80). The use of improvement points creates a fair assessment setting that emphasizes the improvements and efforts and indeed provides every student and group with equal opportunities for success. However, this relies on every student being individually tested and scored each time after group work, which is only feasible for small classes with adequate teaching time.

It is worth noting that a very important element of CL group assessment which contributes to more effective learning is the integration of peer and self-assessment. Traditionally, “Assessment is the principal mechanism whereby staff exercise power and control over students” (Boud et al., 2001, p. 70) whereas students are solely passive recipients of assessment. Modern educationists have been widely aware that when students realize they are not only learners, but also controllers and assessors of their own learning, their sense of ownership of learning is substantially strengthened and moreover their intrinsic learning motivation is greatly improved (Boud et al., 2001; Jacobs & Goh, 2007; Johnson et al., 1998; Slavin, 1995; Wilhelm, 2006). In addition, the use of peer assessment considerably increases the quantity and quality of overall assessment. Through peer assessment, a good variety of different perspectives are likely to be generated, compared with the situation where the teacher is the only assessor of learning. Some studies have found that peers tend to provide each other with more immediate detailed feedback and assessing comments, which are reciprocal for both parties either giving and receiving assessment since these feedback and comments are conducive to critical self-reflection on the part of student assessors and corresponding remediation on the part of student assesseees (Boud et al., 2001; Jacobs & Goh, 2007; Johnson et al., 1998; Wee & Jacobs, 2006). Some researchers (e.g. Holt, 1993; Johnson et al., 1998; Kagan, 1994; McCafferty et al., 2006; Reid, 1993) believe that students can learn as

much, if not more, from their peers as they do from teachers. Therefore, it is essential that results of peer and self assessment should be valued and included as an integral part into the formal assessment of the course. This can contribute to a more accurate measure for learning outcomes, stimulate students to take peer assessing procedures seriously, enhance learners' intrinsic motivation, and facilitate a more active and productive learning environment.

Teacher's Roles in Cooperative Learning

Teachers play a very different role in the CL classroom in contrast to the traditional classroom where they are considered the transmitter of knowledge or "a sage on the stage". The fundamental change CL teachers should make in their role lies in their transfer to a facilitator of learning or "a guide on the side" (Johnson et al., 1998, p. 2:2). Playing a facilitative role involves delegating authority to students and empowering learning so that students are able to make decisions and be responsible for their own learning. Baloché (1998, p. iii) defines empowered learners as "learners who are capable of—and committed to—high levels of meaningful cooperative inquiry, high levels of independent thought, and active and productive participation in a diverse, democratic society". However, on the other hand, delegating authority does not mean that teachers are to be less active but actually to play a even more active and demanding role in the CL classroom (Cohen et al., 1994; Jacobs, 2006; Jacobs & Goh, 2007).

As a facilitator, "teachers play an essential role in helping groups function well" (Jacobs, 2006, p. 38) through a series of procedures of a cooperative lesson. These procedures encompass making pre-instructional decisions (e.g. setting learning objectives, arranging learning materials, grouping students and assigning them individual roles), explaining team tasks and cooperative methods or structures to be used, monitoring and making necessary intervention while students are working in groups, evaluating and processing the quality and quantity of group work together with students, and reflecting on how they have been doing as a facilitator (Brown & Thomson, 2000; Johnson et al., 1998). In addition, more often than not, teachers also need to design cooperative tasks and select or modify CL methods or techniques according to their specific teaching situations before

getting to the stage of explaining team tasks and cooperative methods. This is particularly important when a CL curriculum is not available or the existing CL curriculum is not adequate.

Obviously, facilitating students' learning in the CL classroom means the teacher must be competent in playing multiple roles, which are substantially more challenging than simply passing on information or knowledge to students. Synthesizing CL teachers' roles posited by CL leading researchers (e.g. Baloché, 1998; Gillies, 2007; Holt, 1993; Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994; Sharan, 1994; Slavin, 1995), suggests some basic roles that teachers, as facilitators, should play in the routine process of CL lessons, although not necessarily involving all the roles in a particular lesson.

First, they are controller and instructors. Delegating authority does not mean that teachers are asked to give up control of the class but to exercise control so that cooperative student groups can function well (Cohen et al., 1994; Jacobs, 2006), and "teachers are still active in the usual ways, some of the time—standing in front of the class to explain and demonstrate" (Jacobs & Goh, 2007, p. 30). Actually, there is a consensus that giving instructions on learning content and teaching necessary social skills are important parts of cooperative lessons (Brown & Thomson, 2000; Dishon & O'Leary, 1994; Gillies, 2007; Jacobs & Goh, 2007; Slavin, 1994, 1995).

Second, they are technique selectors, method modifiers and task designers. As mentioned in the section on selecting CL methods and techniques, teachers need to select suitable techniques or modify the existing methods so that the employed techniques or methods fit best in to their particular teaching settings. Along with technique or method selection and modification, another demanding job for the teacher is to design the CL task which "must be set in a way that it engages the entire group" (Brown & Thomson, 2000) and suit students' current academic level and personal interest (Jacobs & Goh, 2007).

Third, they are organizers, guides and encouragers. They plan and organize cooperative lessons by explaining learning objectives, team tasks, individual accountability, and criteria for group success and so on. They guide group work on the side as participants, advisors and encouragers. When students are working on group tasks or projects, one primary function the teacher bears is to encourage and stimulate supportive peer interaction because numerous studies

indicate that the more students interact on their tasks the better they learn (Cohen et al., 1994; Gillies, 2007, Kagan, 1994, Jacobs & Goh, 2007).

Fourth, they are observers, monitors, and interveners. Observing and monitoring student groups serves as a means of knowing what students are doing about their work and how well groups are functioning. This is also an opportunity for teachers to intervene and give extra help when needed to improve task work and teamwork. However, many researchers suggest that “giving students space to solve their own problems is also very important” for learner autonomy and life-long learning, and teachers should “resist the temptation to help students the moment they have difficulty, because by intervening, we deprive students of opportunities to learn from each other and to learn from their own failures” (Jacobs & Goh, 2007, p. 32).

Last, they are assessor and reflectors. As mentioned above in the section on assessing CL group work, teachers work with students to assess and evaluate student performance and achievement by giving constructive suggestions and feedback on how to improve their future team tasks and team cooperation. In the same vein as students processing their group work, teachers should also reflect on their work and performance in facilitating students’ learning (Dishon & O’Leary, 1994; Johnson et al., 1998; Kagan, 1994). Through self-reflection, which is also a very important procedure of teacher action research, teachers have a clear picture of what should be used more often, avoided or improved; accordingly, they set new personal goals and implement action plans.

Cultural Appropriateness of Cooperative Learning in China

Modern technological advances and globalization enable many less-developed nations to get access to the updated approaches from the other side of the world and to “catch up with the most recent innovations that were initiated thousands of kilometres away, thereby taking a huge developmental leap without implementing a research or testing phase” (Phuong-Mai et al., 2006, p. 2). This disregards the impacts of cultural factors on education and therefore may result in false

universalism which assumes that one size can fit all. Since the 1990s, culturally appropriate pedagogies have been widely proposed, which represents an attempt to take cultural factors and differences into account when adopting an educational approach rather than to import educational theories and practices across cultures straight away without consideration of the particular host cultural heritage. Recently, many researchers (e.g., Chen & Hird, 2006; Flowerdew, 1998; Holliday, 1994; Jacobs & McCafferty, 2006; Kramsch & Sullivan, 1996; Li & Campbell, 2008; Liang, 2004; McGroarty, 1993; Phuong-Mai et al., 2006) contend that cultural dispositions toward learning can be different and therefore the way that CL may interface with a particular culture is worthy of consideration.

When studying the general features of a particular culture, Hofstede (1986, 2001) summarized the four dimensions where culture differs: power distance, individualism-collectivism, masculinity-femininity and uncertainty avoidance. Examining these four dimensions of a particular culture is conducive to judging the appropriateness of cross-culture educational methodology in a particular pedagogical context (Jacobs & Ratmanida, 1996; Phuong-Mai et al., 2006).

“Power distance is the extent to which less powerful members of institutions expect and accept that power is distributed unequally” (Phuong-Mai et al., 2006, p. 4). In the case of educational settings, a high level of power distance is associated with the assumption that the teacher is the authority and fount of all knowledge, while students are passive, obedient, and are expected simply to follow. Thus students are unlikely to “see themselves as having valuable knowledge to share with one another and being capable of taking initiative and learning together” (Jacobs & Ratmanida, 1996, p. 106). Hofstede’s second dimension is individualism-collectivism, which concerns how people interact with each other. It is assumed that students from high collectivism societies would be more willing to share with one another and prefer to work collaboratively in groups. In the third dimension, masculinity is defined as being materialist, assertive and self-centred while feminine cultures place value on social solidarity, interdependence and service to others. It is hypothesized that students in feminine societies are inclined to support each other and enjoy group work. The last dimension, uncertainty avoidance, relates to the extent

to which people of a culture feel threatened by uncertainty or unknown situations. Strong uncertainty avoidance implies that students prefer direct instruction from the teacher rather than group work involving self-regulation and learning autonomy.

It is commonly believed that many Asian countries influenced by Confucianism are high along the dimensions of power distance, collectivism, femininity, and uncertainty avoidance, while most Western cultures appear to be the opposite. These stereotypes reveal that both Asian and Western cultures “appear to have traits which both support and oppose the use of groups” (Jacobs & Ratmanida, 1996, p. 107). In view to this situation, some researchers recommend using structured group work, which involves detailed instruction, precise objectives, well-defined tasks and clearly-specified assignments (Phuong-Mai et al., 2006); In other words, this type of group work provides a rigorous framework and steps for students to follow when completing group tasks and interacting with each other, which is just what CL incorporates. On the problems concerning the ELT in CHC countries (e.g. low motivation to use the target language, significant variation in proficiency levels, and large-class teaching), Jacobs and Ratmanida (1996, p. 103) pointed out that “the literature on cooperative learning and task-based language teaching may provide insights into methods of increasing the effectiveness of group activities”, while simultaneously teachers should “use their own local knowledge to adapt group methods to fit their particular contexts”.

China, as the homeland of Confucius, has been long influenced by Confucianism, which is underpinned by three major principles: humanism, faithfulness and propriety. These principles emphasize a variety of traditional values which may significantly affect the implementation and results of CL. Humanism and faithfulness are considered conducive to CL because they highlight individuals’ uniqueness, positive self-concept and self-esteem, awareness of and sensitivity to other, supportive interpersonal relationships, collectivist orientation, cooperation and group cohesion. Thus CL techniques have been strongly recommended for learners from CHC. A good number of studies indicate that learners from CHC contexts (e.g. Japan, Korea, Singapore, Vietnam, Malaysia, Taiwan, Hong Kong, China) prefer working in groups (Deng, 2007; Littlewood, 2000; Le, 2006; Sachs, Candlin, Rose & Shum, 2003; Teng et al., 2004) and perform better in groups (Chen, 2005; Joritz-Nakagawa, 2006; MuGuire, 1992; Park, 2002; Wee & Jacobs,

2006) even in an exam-oriented and highly competitive environment (Flowerdew, 1998). A recent study conducted in China demonstrated similar results: tertiary students expressed a preference for teaching styles that allow peer interaction and collaboration rather than the traditional teacher-fronted direct instruction that involves few student-student interactions (Zhang, 2006). Chen and Hird (2006) also reported that Chinese tertiary students successfully demonstrate a variety of collaborative skills when working together towards a mutual goal.

Propriety is another noteworthy principle, which is deep rooted in CHC contexts. It is based on the premise that society is hierarchically ordered and that societal stability depends upon unequal relationships in terms of age, seniority, rank and authority. This principle breeds two prevailing values: concept of “face” and self-effacement. The concept of “face” demands respect and empathy for others, and that due respect or “face” must be given to each other. It eschews the practice and behaviours which make each other lose “face” by giving or getting overt and public criticism. Self-effacement emphasizes the importance of modesty and “requires that individuals maintain a certain level of humility in accordance with their rank, and do not elevate themselves above others” (Flowerdew, 1998, p. 326). These two values—the concept of “face” and self-effacement—may have both a positive and negative impact on the use of CL.

On the one hand, the concept of “face” consolidates students’ awareness of empathy and concern for others, and also stimulates students to work hard and do a good job so that they, either as individuals or as a group whole, will not lose face before others. Similarly, the concept of self-effacement facilitates group harmony and cohesion, greatly reduces unnecessary conflicts among group mates, and thus enhances cooperation and effectiveness of group learning. But on the other hand, these two values are also very likely to cause pseudo cooperation in the guise of surface harmony. This is because students tend to be passive and reticent in learning, and feel reluctant to challenge while they show their own self-effacement and avoid the issue of “face”. In other words, under the influence of the concepts of face and self-effacement, students are likely to display unquestioning acceptance of what the teacher imparts and avoid giving individual critical comments on peers’ work (Jacobs & McCafferty, 2006; Liu, 2002; Peng, 2007; Phuong-Mai et al., 2006; Tjosvold, Nibler & Wan, 2001). Therefore, these two values may result in rote learning, lack

of autonomous learning and critical thinking, and absence of interaction, which may impair learning.

However, contrary to the common assumption that CHC students lack conflict-resolution skills, some recent studies have found opposing evidence among Chinese tertiary students who exhibit satisfactory skills in proposing opposing views, challenging others' position and resolving conflicts in group work (Chen & Hird, 2006; Tjosvold & Fang, 2004). Chen and Hird (2006, p. 79) found a range of cooperative skills used by students in the foreign language classroom, including not only "assisting each other with lexis, accumulating and modifying ideas", but also "showing agreement and disagreement, and reflecting on how disagreements were expressed in English with different audiences". Tjosvold and Fang (2004, p. 81) found that "Chinese people not only can manage their conflicts openly but they can do so productively and enjoyably". In addition, Littlewood (2000) compared students between Asian and European countries through a large-scale survey relating to their learning attitude and preference for learning styles. He found responses of students from different countries were very similar and Asian students, like their European counterparts, preferred to be active and independent in learning instead of being spoon-fed with facts and textbook knowledge. His survey also revealed that Asian students would like to work cooperatively in groups towards common goals in a friendly and supportive learning atmosphere. He further argued that

if Asian students do indeed adopt the passive classroom attitudes that are often claimed, this is more likely to be a consequence of the educational contexts that have been or are now provided for them, than of any inherent dispositions of the students themselves (Littlewood, 2000, p. 33).

In some sense, Asian students' unwillingness to communicate and passivity in learning should be attributed to the competitiveness and negative interdependence among students in the classroom, which means that one's failure increases the probability others' success. Reflecting its impact on the language classroom, Kagan and McGroarty (1993, p. 65) state that "Would an intelligent person be willing to experiment with new and unfamiliar language forms in a room full of others who, rather than help, hope for their failure? An adaptive response in that situation is to remain

mute”.

Also, there is the possibility that, increasing globalization and cross-cultural exchange results in increasing assimilation of ideas from other cultures. This means that young generations of CHC nations can be considerably influenced by western cultures and thus the cultural gap between the East and the West may gradually be narrowed in the future.

Links between Cooperative Learning and Second Language/Foreign

Language Teaching

It is believed that CL can provide an effective means of developing students' language proficiency through increasing the use of interactive tasks and teamwork among highly-motivated learners in a well-structured supportive learning environment (Richards & Rodgers, 2001). The CL approach and principles of second language (L2)/foreign language (FL) teaching are closely linked with each other in the way that CL activities serve as good solutions to some critical issues in L2/FL teaching and learning (Dörnyei, 1997; High, 1993; Holt, 1993; Kagan, 1994; McCafferty et al., 2006; Sharan & Shachar, 1988). These critical issues will be dealt with in detail when discussion are presented regarding theoretical perspectives and concepts from the literature on L2/FL teaching and learning, which include the input hypothesis, the interaction hypothesis, the output hypothesis, affective factors of anxiety and motivation, and goals of L2/FL teaching.

Input Hypothesis

Successful second language learning must involve enough input, intake and output (Ellis, 1994; Hedge, 2000). Krashen (1982) proposes an input hypothesis, which emphasizes that an important condition for language acquisition to occur is that the acquirer should be exposed to a comprehensible and sufficient input of “i+1” level. If a student's current level is “i”, the input

should contain “i+1”, which refers to the next stage or level. In other words, input should be neither so far beyond students’ reach that they are overwhelmed and thus discouraged, nor so close to their current stage that they are not challenged at all. In practice, direct instruction by the teacher can only provide one dimension of “i+1” level input, which may suit one group of students who are usually average students (with both high and low achievers’ needs ignored). Thus more teaching strategies are called on to provide a greater range of “i+1” level input so that the needs of more students can be satisfied. In this case, CL serves as a good solution through promotive peer interaction generated by structured teamwork; Kagan and McGroarty (1993, p. 64) state that “For all groups, input is made comprehensible through the negotiation process inherent in cooperative learning”.

Interaction Hypothesis

The interaction hypothesis (Long, 1996; Long & Porter, 1985) recognizes the importance of comprehensible input, but also stresses the role of social interaction and meaning negotiation in processing language input and developing language proficiency. In the process of the interaction, both input and output are involved between the participants who are listeners and speakers by turns. Negotiation of meaning occurs when problems for understanding arise and thus listeners ask for repetition or clarification while speakers check for understanding. It is held that simplification, modification, and contextual and extralinguistic clues occurring in peer interactions are conducive to comprehending input and obtaining intake. Peer interaction can result in more comprehensible input and language intake than teacher talk or interaction with native speakers, owing to their similar interests, cognitive levels and language proficiency (Ellis, 1999; Long, 1983; Long & Porter, 1985; Varonis & Gass, 1985). Apart from through providing comprehensible input, meaning negotiation can contribute to acquisition through receiving negative feedback from others by means of recast, and through the opportunities for learners to reformulate their own erroneous utterances in a more target-like way (Long, 1996).

In the context of FL learning, the position of peer interaction turns out to be more valuable

because it compensates for the lack of authenticity in both language input and output through creating real-life communication in the classroom (Ellis, 2009b). However, in traditional classrooms where “teachers control the discourse”, learners are usually “reluctant to signal their lack of comprehension and to negotiate understanding, preferring instead to either wait and see if they can work things out later or, alternatively, to abandon any attempt to comprehend” (Ellis, 1999, p. 223). It is argued that “giving learners control of the discourse is one way of making the classroom acquisition-rich” (Ellis, 1999, p. 219).

Likewise, Hedge (2000, p. 13) asserts that “There is a principle underlying current ELT practice that interaction pushes learners to produce more accurate and appropriate language, which itself provides input for other students”; consequently, group work, which allows more student discourse control, has become a common feature of contemporary classrooms. However, some studies (e.g. Pica & Doughty, 1985) have found that group work generates little negotiation of meaning in the language classroom when there is domination by a few individual students who are usually high-achievers. This led the researchers to pose the question of what methods are best for active involvement of all group members in peer interactions; for this question, CL can probably provide a satisfying answer with its major principles such as positive interdependence, individual accountability, promotive interaction and equal participation.

Output Hypothesis

Swain (1985; 1993) proposes the output hypothesis, which suggests that comprehensible input is not a sufficient condition for language acquisition, and that language production is also a must-do for effective language learning. In order for learners to improve their L2/FL proficiency, they must produce the target language via speaking or writing and receive feedback on the comprehensibility of their output. Getting feedback from the teacher and peers enables learners to test hypotheses and improve their knowledge of the language system. Lack of opportunities for output production forms a primary reason for inadequacy in real communication with learners who have already demonstrated a good command of receptive skills (e.g. reading skills as well as grasp of grammar

rules and vocabulary). Output production brings some benefits that input alone cannot, for instance, promoting fluency through using language meaningfully, motivating learners to engage in syntactic processing of language, and offering opportunities for feedback from others and hypothesis testing as to how language works in terms of appropriateness, correctness and comprehensibility (Swain, 1993). Use of group work in the language classroom not only increases the amount of output production but also enables students to perform a greater range of language functions and is likely to result in greater complexity of language output (Ellis, 1994, 1999, 2009b; Hedge, 2000; Kagan & McGroarty, 1993; Long & Porter, 1985; Pica & Doughty, 1985). Moreover, Magee and Jacobs (2001) compared the use of three teaching modes—teacher-fronted instruction, unstructured group work and CL—with tertiary international students who learned the Chinese language in Singapore. They found that the students produced significantly more output and took more turns in the two types of group work than in the teacher-fronted mode, and significantly more in structured CL group than in the unstructured group. There is concern that language production generated among peers may contain more errors, for teachers' direct intervention and supervision are not much involved. However, this has not been found to be the case (Ellis, 1994; Jacobs & McCafferty, 2006; Pica & Doughty, 1985); that is, relevant research has showed no substantial difference in the number of output errors between the students exposed to group work and those with teacher-fronted instruction.

Affective Factors

Many affective factors are held to be critical in L2/FL teaching. Krashen (1982) has proposed the hypothesis of affective filter, which calls attention to language learners' affective domain. The hypothesis suggests that high anxiety, low motivation and negative attitude towards learning may raise the affective filter and form a mental block that inhibits students' comprehending language input or prevents comprehensible input from being used for acquisition. Thus in order to improve the outcomes of language teaching, attention should be paid to reducing debilitating anxiety and increasing learning motivation. Many studies link anxiety in language learning to competitiveness

typically involved in traditional whole-class teaching, where students compete to outdo others in tests or to gain the teachers' approval (Ellis, 1994; Hedge, 2000; Kagan, 1994). When the competition is too overpowering, students with excessive anxiety are likely to retreat into complete passivity or even withdraw from the learning experience. Consequently, Chamot and O'Malley (1994) recommend cooperation within groups as an important social/affective strategy to alleviate excessive anxiety and develop positive learning attitudes in L2/FL language classrooms.

Motivation is another important element in the affective domain, which has high positive correlations with successful learning (Ellis, 1994; Harmer, 1998; 2007; Hedge, 2000). In general, motivation can be categorized into two types: intrinsic motivation (which derives from the learner's personal interests and inner needs) and extrinsic motivation (which comes from external sources such as passing examinations and getting material rewards). Extrinsic motivation is integrated in CL when group members are interdependent to complete a group goal for group rewards or recognition. It is assumed that extrinsic motivation is an effective and practical means of achieving successful learning in the FL context (Dörnyei, 2006), where few students feel internally motivated especially at the initial stage of learning. Dishon & O'Leary (1994) and Johnson et al. (1998) hold that having external reasons to work together in cooperative groups can lead to the development of intrinsic motivation to learn in the long run. Jacobs and Goh (2007, p. 38) contend that CL teamwork "can provide a means of promoting intrinsic motivation, in which students strive to meet internally-set goals". An effective way to enhance students' motivation is through building up group cohesiveness and gaining positive learning experiences, which is inherent in the CL activities (Clement, Dörnyei & Noels, 1994; Dörnyei, 1997; Hedge, 2000; Jacob & Goh, 2007; Kagan, 1994, Noels, et al., 2000). This is because CL principles emphasize the importance of creating a supportive, caring and positively-interdependent environment, where students feel safe to speak and highly motivated to contribute to the group goal. According to Sharan and Shachar (1988), cooperative groups create a more congenial and accepting learning environment in which students at the disadvantaged end of spectrum, in terms of either academic level, language ability or social status, can participate more freely and make fuller use of their

verbal abilities than in the traditional classroom. Cohen et al. (1994) and her colleagues also report that well structured CL tasks are conducive to L2 learning by stimulating the students' needs for communication, maximising peer interaction and generating comfortable learning environments. The "cohesiveness-performance effect in cooperative learning can be particularly strong in language classes because learners' communicative skills are developed primarily through participatory experience and verbal interaction in real world language tasks" (Dörnyei, 1997, p.485).

Goals of Second Language/Foreign Language Teaching

A review of the perspectives of researchers and practitioners (e.g. Ellis, 1994, 1999; Harmer, 2007; Hedge, 2000; Jacobs & McCaferty, 2006) in the L2/FL field suggests that L2/FL teaching is aimed at two primary goals: development of communicative competence and cultivation of learner autonomy. This is also reflected in the *College English Curriculum Requirements*, which states that the objective of College English teaching involves developing students' ability to use English in real-life communication and enhancing their ability to study independently (MOE, 2007).

Development of Communicative Competence

Language is a system used for human communication. As Hedge (2000, p. 44) states "The ability to communicate effectively in English is now a well-established goal in ELT". Professionals in the modern L2/FL teaching have reached a consensus that the ultimate goal of L2/FL teaching lies in the development of learners' communicative competence. As for the key components of communicative competence, Hedge (2000, p. 46-55) summarizes them into five categories: linguistic competence, pragmatic competence, discourse competence, strategic competence and fluency.

Linguistic competence is related to knowledge of the language itself, including the aspects of "spelling, pronunciation, vocabulary, word formation, grammatical structure, sentence structure,

and linguistic semantics” (Hedge, 2000, p. 47). Pragmatic competence is concerned with two types of ability: illocutionary (i.e. knowing how to use language to successfully perform a function or achieve an intention) and sociolinguistic (knowing how to achieve verbal or non-verbal communication appropriate for a particular social context). “Part of communicative competence in a foreign language is knowing what is appropriate, what is incongruous, and what might cause offence” (Hedge, 2000, p. 50). Discourse competence, also termed as textual competence, involves the ability to use discourse markers or cohesive devices to create coherent written or spoken discourse. Strategic competence refers to the ability of speakers to use verbal or non-verbal communication strategies to compensate for breakdowns in communication or to improve the effectiveness of communication. Fluency, normally relating to oral language production, is defined as “the ability to link units of speech together with facility and without strain or inappropriate slowness, or undue hesitation” (Hedge, 2000, p. 54).

Clearly, these five dimensions of competence, especially the latter four, cannot be achieved solely through traditional teacher-fronted instruction which focuses on the transmission of knowledge of language rather than the ability to use the language. As a result, contemporary L2/FL literature stresses that students should be provided with plenty of opportunity to practise the target language and communicative activities should be integrated as part of classroom procedures. The key element for successful use of CL in L2/FL is to design and use cooperative structures and activities, which involve a wide range of language functions and social interactions aiming at maximizing active participation and authentic verbal communication in the classroom (High, 1993; Jacobs & Goh, 2007; Kagan, 1994; Richards & Rogers, 2001).

In addition, one commonly accepted way to enhance fluency in L2/FL is to teach formulaic expressions for particular functions or purposes (Ellis, 1994; Harmer, 1998; 2007; High, 1993; Hedge, 2000). Formulaic expressions are items of prefabricated language, which are learned holistically as lexical chunks that may take the form of phrases, clauses and sentences. It is assumed that once taught, these expressions are likely to be retrieved quickly from memory as a whole, and will facilitate learners’ fluency. Numerous CL researchers (e.g. High, 1993; Jacobs & Goh, 2007; Kagan, 1994) advocate that, along with assigning students CL group tasks, they should

get access to formulaic expressions in the target language which allow them to function socially and achieve particular communication goals. According to Ellis (1999, p. 223), “providing learners with the formulae they need to carry out negotiation (e.g. a list of ways of requesting clarification)” serves a good way of reducing learners’ reluctance to communicate. There is empirical evidence that allowing learners (especially those of low or intermediate proficiency) time to plan “what content to express” and “what language to use” results in greater fluency (Ellis, 2009c, p. 474).

Cultivation of Learner Autonomy

Another important objective for L2/FL teaching is to cultivate learner autonomy. A review of L2/FL literature has found a variety of synonyms for learner autonomy, such as autonomous learning, independent learning, and self-regulated or self-reliant learning. These concepts all highlight the significance of enhancing students’ ability and inclinations to take responsibility for their own learning (Ellis, 1999; Harmer, 2007; Hedge, 2000; Jacobs & McCafferty, 2006). “This is because language is always too complex and varied for there to be enough time for students to learn all they need to in a classroom” (Harmer, 2007, p. 394) and they need to develop into autonomous life-long learners. However, successful cultivation of learner autonomy depends substantially on what kind of classroom teaching procedures students are exposed to. It is universally agreed in the L2/FL educational circles that the traditional teacher-dominated classroom can only create passive and teacher-dependent students, rather than active, independent and self-directed learners. Many researchers (Ellis, 1994, 1999; Harmer, 2007) argue for the critical role of group work in enabling students to have control over their own learning process. A traditional language classroom “where there is very little opportunity for learner discourse control, particularly for learners who have progressed beyond the beginner stages of acquisition, may impede the development of communicative proficiency” (Ellis, 1999, p. 228). In the CL classroom, teachers as facilitators delegate authority to students and empower learning by allowing them to plan, manage and assess their teamwork. In this way, students have more freedom and

accountability for their own learning (Baloche, 1998; Jacobs & MaCaffery, 2006; Kagan, 1994). In addition, CL generates positive learning attitudes and successful learning experiences, which form a critical element for sustaining life-long learning.

There is a concern that “ideas about independent learning are not so easily applicable to Asian cultures” because Asian students are assumed to be more passive and spoon-fed and this may form an obstacle to self-directed learning which has its origin in western cultures (Hedge, 2000, p. 100). However, Hedge also points out that it does not mean learner autonomy goes against Asian educational philosophy, for approximately 2,500 years ago, Confucius stated, “Giving a man a fish feeds him for a day; teaching a man to fish feeds him for a lifetime”, which has been much quoted in the L2/FL literature on learner autonomy and learning strategies. Researchers (e.g. Joritz-Nakagawa, 2006; Littlewood, 2000; Wee & Jacobs, 2006; Zhang, 2006) found that passive learning and spoon-feeding are not inherent in Asian students; on the contrary, they prefer to exhibit control over and take responsibility for their learning. Some studies also show that collaborative group work facilitates the development of autonomy with Chinese FL learners despite their previous experience of teacher-directed instruction in the classroom (Kell & Newton, 1997). However, it is assumed that students would be well-supported with relevant strategy training and instruction through the early stages of fostering learner autonomy, especially in the context where the educational tradition has not offered students this experience (Hedge, 2000). This assumption echoes the insight into the importance of teaching relevant cooperative skills that are essential for teamwork but not familiar to students at the initial stage.

Links between Cooperative Learning and Tertiary Students

Tertiary education is aimed at developing students’ knowledge and skills not only in professional but also social domains so that they can meet the requirements of society and make useful contributions to their professions in the future. According to Kagan (1994), the most frequent reason for a person to be fired from his or her first job is not lack of professional

knowledge or skills, but rather lack of cooperation and interpersonal skills. It is also true that individuals who have made remarkable achievements typically attribute their success to cooperative efforts. Undoubtedly, professions in the future will involve more intensive divisions of labour and exchanges of information, which demand cooperation and communication skills (Kagan, 1994; McWhaw, Schnackenberg, Sclater & Abrami, 2003). In view of this, tertiary students must get adequate cooperative experiences at college and understand the true value of cooperation and teamwork.

The research on CL in relation to tertiary education mainly started in the 1990's when educationists increasingly realized the importance of cooperation to tertiary students, especially for their academic accomplishments and future professional development (Riordan et al., 1997). Since then, there have been some studies on CL with tertiary students in quite a few areas such as social sciences, accounting, laboratory chemistry, physics, business communication (Riordan et al., 1997), management, law, information technology and engineering (Boud et al., 2001), the majority of which reveal a positive effect on students' cognitive, affective and social domains. However, compared with the ample evidence that supports the use of CL in grades 2-9, there are still relatively few studies examining the effectiveness in post-secondary education. Indeed, one controversy about CL is on the issue of whether CL is effective at all levels (Slavin, 1997), so more research is called for to determine the appropriateness and effectiveness of CL at the tertiary level (McCafferty et al., 2006; McWhaw et al., 2003; Richards & Rodgers, 2001; Slavin, 1997).

Some researchers question the applicability of CL to tertiary students because it tends to be highly structured, rigorously prescriptive and more directive about how to work together, which they believe may impede students' development in higher-order thinking skills, initiative and autonomy in learning (Bruffee, 1993, 1995; Matthews, Cooper, Davidson & Hawkes, 1995; Oxford, 1997). Consequently, these researchers assume that CL is more beneficial for students at elementary and secondary school levels, while less-structured collaborative learning, which allows for more power and freedom over the learning process and assessment, is considered more suitable to postsecondary or tertiary levels for non-fundamental knowledge and high-order thinking. On the other hand, some other researchers deny there is a clear distinction between CL and

collaborative learning because of the considerable overlap focusing on positive peer-to-peer interaction and fulfilment of mutual goals (Boud et al., 2001; McCafferty et al., 2006; Romney, 1997).

For the purpose of this research, the author prefers to take CL and collaborative learning as synonyms and think of CL as equally suitable for College English learners. First, some CL structures (e.g. GI and Co-op Co-op) provide plenty of opportunities for students to make decisions and have control over their learning, and allow students a great deal of development in their critical thinking and problem solving skills (Brown & Thomson, 2000; Gillies & Ashman, 2003b; Kagan, 1994; Sharan & Shachar, 1988; Sharan & Sharan, 1994). Second, there is a strong influence from traditional direct instruction to which Chinese tertiary students have been exposed for more than ten years, and consequently they tend to be more passive and less self-directed in learning. Thus well structured and clearly oriented activities and lesson designs, are more suitable for Chinese College English learners, especially in the initial stage, so that they know where to go and what to do. In this way, mutual participation and adequate interactions can be greatly facilitated.

In addition, tertiary students are adults aged over 18 years, who learn differently from children. Adult students learn best when they can get access to opportunities for practising and experiencing what they are learning about, when they can take responsibility for their learning and feel their emotional well-being has been taken into account (Gibbs & Habeshaw, 1989). Vella (2002) holds that teaching adults must take into consideration the role of interaction between the teacher and learners, and claims that the use of cooperative peer interaction is probably the most powerful learning strategy in adult learning. She also identifies some critical principles which must be observed in adult education, such as safe learning contexts, sound relationships between teacher and learner and among learners, learning by doing, respect for learners as active participants and decision makers, and use of teamwork. Obviously, these principles are closely associated with CL; in other words, CL has a great potential to serve as an excellent means for carrying out these principles and achieving the expected outcomes.

Links between Cooperative Learning and Large-Class Teaching

In Western cultures, a class of over 35 students is considered large and difficult to teach (Bennett, 1996). Sharan (2003) believes that a precise definition of a large class should take into consideration not only the number of students but also the number of teachers engaged in teaching the class, the instructional methods, subject matter and the age of students. Chinese College English classes, with over 50 students on average and only one teacher available per class, are definitely categorised as large.

It is often taken for granted that large-class teaching can only take the form of teacher-fronted lecturing. However, lecturing might be as effective as other methods in providing and transmitting information, but definitely has drawbacks in teaching language skills, which must involve active participation and interaction with each other (Brown & Atkins, 1991; Cannon & Newble, 2000; Gibbs & Habeshaw, 1989). A considerable amount of research indicates that there are some problems related to teacher-centred large-class lecturing, and it is recommended that CL serve as an avenue for coping with them (Brown & Atkins, 1991; Cannon & Newble, 2000; Gibbs & Habeshaw, 1989; High, 1993; Kagan, 1994; Kagan & Kagan, 1994; Sharan, 1994, 2003; Slavin, 1995).

Firstly, students' short attention spans and frequent disruptive behaviours are common complaints from teachers and lecturers who instruct large classes. "15 minutes into a lecture learners will be performing much less well than at the start" (Gibbs & Habeshaw, 1989, p. 30). This problem primarily lies in the fact that very few teachers have the capacity to present long lectures in a strikingly dramatic and attractive way and students tend to lose attention quickly during a passive and boring learning task. However, this decline in attention can be remedied if we bring in some cooperative small group activities for a change. CL, which involves more active participation, peer interaction and personal relevance, contributes a lot to refreshing learners from passive learning, restoring their learning performance, attracting their attention and extending their on-task time (Brown & Atkins, 1991; Cannon & Newble, 2000; Sharan, 2003).

Secondly, teacher-fronted large-class lecturing does not encourage mutual interaction and communication. For one thing, there are a large number of students, few of whom can be called upon to speak because of the sequential structure where only one student is allowed to speak each time. For another, even if occasionally students get the chance, they mostly feel reluctant to put themselves in the spotlight by asking or answering a question, or engaging in any kind of interaction, which is especially typical in Eastern cultures (Cannon & Newble, 2000). The simultaneous structure and supportive learning context incorporated in CL are an excellent means of sorting out these problems (High, 1993; Kagan, 1994; Kagan & Kagan, 1994, 2009; McCafferty et al., 2006; Stone, 1994).

Thirdly, large classes probably involve wide differences in a variety of dimensions such as academic levels, interpersonal skills, personal interests and personalities. This intensifies the difficulty of teaching the students, and makes it impossible to adjust the learning materials and instructional methods to everyone's level and taste. As a result, most teachers and lecturers simply try to meet the demands of those assumed to be average, ignoring those at both ends who are high and low achievers. It is widely believed that CL works well not only in homogeneous groups but also in heterogeneous groups where students are motivated to facilitate each other's learning and, thus diversity within the group is converted into a rich resource rather than a problem (Brown & Thomson, 2000; Johnson et al., 1998; Kagan, 1994; Slavin, 1995).

As for how to achieve successful large-class teaching, Harmer (2007) proposes a number of key elements that teachers should bear in mind. These elements include that: teachers should be more organized with pre-set tasks to conduct lessons; teaching procedures and class management routines should be established with students at the start of a course; strategies such as pair work, group work and peer tutoring should be used with personal responsibility well assigned; and worksheets should be designed for group activities. It is apparent that all these elements are well linked with the principles of CL.

However, the incorporation of CL into large-class teaching does not occur automatically and naturally, but is conditional on some supportive factors: adequate classroom size, teachers' professional knowledge and skills in CL, open organizational policy and assessment system,

collaborative teaching teams, and a supportive educational system (Sharan, 2003). So far CL has been increasingly incorporated into online large classes or learning communities; this type of structure is usually termed computer-supported/online/network-based/web-based collaborative learning, and has turned out to be effective and meaningful in a wide range of disciplines (Roberts, 2004, 2005; Warschauer & Kern, 2000). However, an extensive literature review has found very few studies concerning the impacts of CL with classroom-based large-class teaching, especially in the L2/FL field.

Recent Studies on Using Cooperative Learning in Teaching English as a Second Language/Foreign Language

Although there is ample evidence that supports the use of CL in a variety of educational settings, very few studies are related to L2/FL teaching (Jacobs & Goh, 2007; McCafferty et al., 2006), especially comparing the effectiveness of CL with that of traditional instruction in teaching English as a L2/FL. This section is aimed at presenting recent studies regarding the use of CL in teaching English as a L2/FL, around the world as well as in China.

Experimental Studies in Different Countries

In recent years, a small number of experimental studies have been conducted to compare the effectiveness of the CL approach with traditional whole-class instruction in teaching English as a L2/FL. A thorough database search has found seven studies which were conducted in Israel, Japan, Lebanon, Hong Kong, Thailand, Taiwan, and Turkey. These studies all use pre-test-post-test control group experimental or quasi-experimental designs and explore the impact of CL on a wide range of aspects including listening, speaking, reading, writing, vocabulary, grammar, learning motivation, attitudes towards learning, academic self-esteem and a feeling of school alienation.

In Israel, Bejarano (1987) evaluated the effects of the CL methods of GI and STAD and the whole-class method on academic achievement in three aspects: listening, reading and discrete-points of grammar and vocabulary. The experiment lasted four and a half months with 664 junior high school students as participants. The results indicated that students in the both GI and STAD classes registered significantly greater improvements than their peers in the whole-class setting on the listening comprehension test, but not on the test of reading comprehension. On the discrete-point test of grammar and vocabulary, students in the STAD classes were found to outperform their peers in the GI classes and the traditional setting.

In Japan, McGuire (1992) carried out a three-week study with 87 tertiary learners as participants. The results indicated that students in CL groups wrote significantly longer skits (with a substantively greater number of turns between skit characters) than students working individually in traditional classrooms. As for occurrences of errors in the skit writing, the difference between the two methods also approached the level of significance ($p=0.07$) in favour of CL. As a result, the author argued that CL lessons with clearly defined tasks and roles fit in well with the type of structured learning that Japanese students expected and that CL could be an effective way to improve the EFL teaching and learning in Japan.

In Lebanon, Ghaith (2003) conducted a ten-week comparative study with 56 high-school learners, aiming at investigating the effects of the Learning Together Model on students' reading achievement, academic self-esteem and feelings of school alienation. The results revealed a statistically significant difference in favour of CL on the variable of reading achievement, whereas little difference was found between the two methods on the dependent variable of academic self-esteem and feelings of school alienation. Ghaith (2003, p. 460) explained that "significant gains in academic self-esteem and school psychosocial adjustment are unlikely to be achieved in the course of short experiments and cooperative interventions".

In Hong Kong, Sachs et al. (2003) carried out a one-year comparative study with 30 students in secondary school classrooms, which found no significant differences between the CL approach and teacher-centred methods in improving students' speaking competence. However, this finding needs to be interpreted cautiously because of the limited exposure of students to the intervention in

this research, that is, CL tasks were used on average only once a month for the duration of the research. Sachs et al. (2003, p. 350) stated that “the project does not allow for a robust comparison of traditional and cooperative learning” and that “it remains to be seen whether learners who receive greater exposure to cooperative learning would outperform those who do not”. Very interestingly, although the scores on the oral tests did not show the superiority of CL, the lesson observation showed that students in CL groups “were able to use some English to assist the accomplishment of the tasks and above all were motivated to try” (Sachs et al., 2003, p. 355). In addition, the interview with the students and teachers revealed that the majority of them were very positive about the CL methods and liked to use them more often in the future. The teachers reflected that the main constraints on the frequent use of CL in the educational context of Hong Kong were the tight teaching syllabus and their limited teaching hours. They felt that it was difficult to squeeze in sufficient time to carry out CL tasks in classes and that they needed to have more time and flexibility to carry out their teaching more effectively.

In Thailand, Waugh, Bowering and Chayarathee (2005) compared the impact of CL versus traditional teaching, on 96 grade six Thai students, with an intervention of 16 hours’ duration. Results provided evidence to support the use of CL in improving students’ reading comprehension as well as attitudes and behaviour towards learning English as a second language. However, a critical inadequacy of this study is the absence of a description of the specific CL methods used in the classroom, so the readers cannot gain a clear picture of what particular classroom procedures took place in the experimental teaching.

In Taiwan, Chen (2005) conducted a similar experimental study by using a combination of STAD, TGT, Jigsaw II and GI with 100 tertiary EFL learners for eight weeks of intervention. Results suggested that the CL methods were more effective than the traditional grammar-translation method in enhancing students’ speaking and listening ability as well as intrinsic motivation, but there was little difference between the two methods in teaching reading. In this study, CL was also found to be better in promoting students’ intrinsic motivation. However, similarly to Waugh et al. (2005), the author did not make it clear how the four different CL methods were combined and used in the Taiwanese educational setting.

Gömleksiz (2007) carried out a four-week study with 66 engineering students in a Turkish university, aimed at investigating the differences between the CL method of jigsaw II and whole instruction in improving students' vocabulary knowledge, use of active-passive voice and attitudes towards English learning. Results revealed statistically significant differences in favour of CL in all the three areas. That is, compared with the control group, the experimental CL group demonstrated significant improvements in vocabulary knowledge, accurate use of active-passive voice; meanwhile, CL exhibited a significant positive effect on students' attitudes towards learning English and promoted better interactions among students as well.

So it can be concluded from these studies that CL can produce positive outcomes in one way or the other in spite of the fact that the findings were quite different and inconsistent in certain aspects (e.g., reading, speaking, vocabulary and grammar). This inconsistency, to some extent, can be explained by the assumption that the efficacy of CL can be varied due to different educational contexts, host cultures, duration of the study and specific CL methods used. However, these studies generally used different CL methods in their interventions, and this makes it hard to do a further analysis based on specific methods.

Recent Studies in China

In the field of EFL teaching in China, as the important role of group work is more and more highlighted in the classroom, CL, as a type of group work with many potential educational benefits, has drawn more attention from researchers and teachers since the last decade. A detailed search of the literature has located a number of relevant studies conducted in China, which can generally be classified into three categories according to the research nature.

The first category includes those that simply present discussions at the theoretical level about the possibility of applying CL in ELT in China (e.g. Chen, 2008; He, 2003; Li, Dong & Ma, 2006; X. Li, 2007; Liu, 2007). These authors have reached an agreement that CL is worthy of exploration in China considering the potential benefits it can bring to EFL learning. However, very unsatisfactorily, these articles are mostly a repetition of previous theories and do not give insights

into how to actually apply CL into the EFL classroom in China. What makes the research more inadequate is that some authors (e.g. He, 2003; Li et al., 2006) were not clear about the fundamental principles and elements of CL: they mistook group work involving some peer interactions for CL.

The second category comprises some descriptive studies on preliminary trials of CL in EFL classrooms (e.g. Deng, 2007; Guo, 2004; Luo, 2007; Tang, Zhang & Dong, 2009; Zhang & Zhao, 2004). These studies found that CL activities brought numerous positive outcomes to EFL classrooms, such as in the aspects of language skills, active participation, relations with peers and class attendance. However, it is noteworthy that a misconception of CL also inhibits the quality of the studies in this category. Some researcher (e.g. Luo, 2007; Zhang & Zhao, 2004) did not show their knowledge of the two CL fundamental principles—positive interdependence and individual accountability—in their articles; and others (e.g. Guo, 2004) were aware of the existence of these principles but failed to appropriately integrate them into classroom practices.

A very small number of studies (e.g. J. Li, 2007; R. Li, 2007; Peng & Qi, 2006; Wang, 2004; Zhang, 2007) belong to the third category which are empirical studies based on rigorous pre-test-post-test control group quasi-experimental design. The results of these studies are generally in favour of CL, suggesting that CL methods are superior to traditional whole-class teaching in improving students' English proficiency, learning attitudes and learning environment. However, an examination of their experimental procedures locates some critical weaknesses in these studies. First, some researchers (e.g. Peng & Qi, 2006) simply stated that CL methods were used with experimental groups without providing further information about what specific techniques were involved. Second, in most studies (e.g. J. Li, 2007, R. Li, 2007; Wang, 2004), positive interdependence, individual accountability and other CL principles were not found to be integrated in the designed activities. Thus it is group work (such as group discussion, group learning and checking, pair work, peer editing and role-play), rather than CL techniques, that was used in these studies. As a result, some problems were reported to have emerged, such as lack of group cohesion among students, social loafing and free-riding among group members. Third, no study is concerned with the adaptation of CL to the particular EFL context in China. For instance,

Zhang (2007) imported STAD without any modification into a large class of 60 students, and stated that “During each class, the teacher will choose one part of the exercise which closely relates to the parts they have listened (to) as the quiz” (p. 38); However, Zhang did not give useful information on how the teacher coped with the huge work of marking and calculating improvement points for 60 individual students twice a week. Undoubtedly, this technique of frequent quizzes is very impractical in the Chinese EFL teaching context, which is characterized by a shortage of teaching staff, limited teaching hours and large-size classes.

Therefore, it can be concluded that, although CL is becoming increasingly interesting to EFL teachers and researchers in China, the current studies on CL are far from adequate either in terms of quality and quantity. Encouragingly, these studies have sent a positive message, that is, group work is applicable to Chinese EFL classrooms and likely to bring multi-dimensional benefits. Furthermore, it can be inferred that if the usual group work can fit in the context, CL, as a special type of group work with an emphasis on group cohesion and a supportive learning environment, is more likely to work well with Chinese EFL learners.

The focus of this research is adapting CL methods for the College English teaching in China, and investigating the effects of CL on tertiary EFL learners on three dimensions: English proficiency, learning motivation and social skills.

CHAPTER THREE: METHODOLOGY

Overview

The research questions in this study focus on whether cooperative learning (CL) is better than traditional instruction in developing Chinese EFL learners' English language proficiency, learning motivation and social skills. These research questions require the use of quantitative research methodology in this study. This chapter starts with some key concepts essential for a quantitative study. It is followed by an introduction of participants and the general research procedure, which includes a pilot study and a main study. The details of the intervention procedure are provided, focusing on different teaching methods used with the intervention group and the comparison group. This chapter also provides details of the three measures used in this research: the College English Test (CET), the Language Learning Orientations Scale (LLOS), and the Social Skills Scale for Chinese College English Learners (SSS-CCEL). At the end of the chapter is an explanation of specific techniques used to analyze the data and principles which guided the data analysis.

Use of Quantitative Research Methodology

The selection of research methodology is decided by the type of research questions (i.e. quantitative or qualitative), which indicate the nature of data, the number of participants, and research process (e.g. the way to collect, analyse and display data) (Cohen, Manion & Morrison, 2007; Drew, Hardman & Hosp, 2008; Gay, Mills & Airasian, 2009; Mutch, 2005). A quantitative research question begins with a proposition or theory to be tested, which calls for numerical data from a relatively large number of participants. The questions that this research was intended to answer can be considered a good representation of this type. The hypothesis of this research is that

CL is better than traditional instruction in developing Chinese EFL learners' English language proficiency, social skills and learning motivation. To test this hypothesis, three sets of numerical data are required, which enable quantitative comparison of the effects of the two different methods on the three aspects of student functioning.

Quantitative research should involve the following key features (Drew et al., 2008; Gay et al., 2009; Mutch, 2005):

- 1) Using deductive logic to work from a hypothesis and then to gather data to test it;
- 2) Focusing on many examples that are representative of a population;
- 3) Following a linear and structured research design with variables and categories set in advance;
- 4) Gathering quantitative data through instruments such as tests and questionnaires;
- 5) Using a researcher approach that is detached and objective;
- 6) Analysing data using descriptive and inferential statistics;
- 7) Displaying data as graphs, tables, and diagrams;
- 8) Reporting in a detached, third-person style.

The two most commonly used quantitative methods are surveys and experiments (Cohen et al., 2007; Drew et al., 2008; Gay et al., 2009; Mutch, 2005). A survey is very useful for gathering large-scale data in order to generalize to a population, and the questionnaire is the most useful tool of the survey when there are a large number of participants. An experiment compares variables under controlled conditions, and includes two different designs: the true experiment, and the quasi-experiment (Campbell & Stanley, 1971; Cohen et al., 2007; Gay et al., 2009; Mutch, 2005). "Often in educational research, it is simply not possible for investigators to undertake true experiments" (Cohen et al., 2007, p. 282). This is because there is little chance to achieve random assignation of participants to control and experimental groups, and also it is impossible to fully control extraneous variables associated with human beings in educational settings.

Quasi-experiments, which do not require full randomization, are often referred to as "less controlled versions of experiments" or "compromise designs". However, this does not mean that control is no longer important for this design. Actually, tightly controlling extraneous variables

and minimizing outside influences should be the central focus of experimental researchers, so that they are able to “attribute a performance level, difference, or change to the experimental variable (the treatment)” (Drew et al., 2008, p. 140). Among several forms of quasi-experiments, one which is called the pre-test-post-test non-equivalent control group design is well-recognized and most commonly used in educational research. This design is believed to be comparatively rigorous because it involves the pre-test and post-test which are equivalently administered to both groups and this addresses the weakness that the two groups are not precisely matched. For field settings where matching is impracticable, it is suggested that researchers use samples from the same population or samples that are as alike as possible (Kerlinger, 1970). Likewise, to cope with realistic conditions outside the laboratory, Robson (1993, p. 46) recommended using “the comparison of ‘intact groups’ (e.g. existing classes in schools) rather than samples selected and allocated for the purposes of the study”.

Participants

This research was conducted in a university in the North of China from October 2007 to July 2008. Participants were first-year College English learners from 14 different non-English major subject areas. They had been streamed into level-one classes based on an English proficiency test which all freshmen in the university are required to take before starting their study of College English. Participants were composed of three intact classes which were randomly selected from over 30 level-one classes. They are respectively, a pilot class (N=49), an intervention class (N=52) and a comparison class (N=48) (see Table 1).

Table 1: Demographic information on participants in the 3 groups (N=149)

Group	No	Gender		Age		EFL learning (yrs)	
		Male	Female	Mean	Range	Mean	Range
Pilot	49	46.9% (23)	53.1% (26)	19.10	17-21	7.94	6-13
Intervention	52	46.2% (24)	53.8% (28)	19.60	18-21	8.17	6-14
Comparison	48	64.6% (31)	35.4% (17)	19.56	18-22	7.88	6-14
Total	149	52.3% (78)	47.7% (71)	19.42	17-22	8.00	6-14

The pilot class was composed of 23 males and 26 females. They were aged from 17 to 21 years with a mean age of 19.10 years and had English learning experience ranging from 6 to 13 years with a mean of 7.94 years. There were 24 males and 28 females in the intervention class. They were aged from 18 to 21 years with a mean of 19.60 years and had been studying English for 6 to 14 years with a mean of 8.17 years. In the comparison group there were 31 males and 17 females. Their ages ranged from 18 to 22 years with a mean of 19.56 years and their experience of studying English varied from 6 to 14 years with a mean of 7.88 years.

Level-one classes were composed of learners of either average or low English proficiency because high-achievers had already been selected into higher level classes. In spite of this university-wide screening process, level-one classes typically display a surprising degree of diversity in their English proficiency.

Research Procedure

Approval to conduct the research was obtained from the Human Ethics Committee of the University of Canterbury (Reference No. 2007/109). The researcher was the instructor of the College English course for all participants. She had taught English for over ten years and had a particular interest in the application of CL in tertiary EFL teaching. Prior to the study, she had participated in CL workshops, observed CL classes and familiarized herself with major CL

methods. She also spent several months conducting a pilot study in order to try out some CL techniques with College English learners in China. All these efforts and experiences enabled her to adapt CL methods for College English teaching. The field research lasted two semesters, with 13 weeks in the first semester spent on the pilot study and 18 weeks in the second semester spent on the main study. In each week, students studied English for four hours, which consisted of two hours of reading and writing in the classroom and two hours of listening and speaking in the language laboratory. The teaching content covered in the intervention, during the semester of 18 weeks, is summarized in appendix 1 and appendix 2.

Pilot Study

In the first semester lasting from October 2007 to January 2008, a level-one class was randomly selected as the pilot class. In the first session with the pilot class the researcher presented a brief introduction about the research and the CL approach to be used with them. Then consent forms were distributed for students to sign. Since one student preferred traditional teaching, she was transferred to another class on the same teaching schedule after getting approval from the college administration. This made the pilot group 49 students. The pilot study was aimed at fulfilling two major tasks: piloting two questionnaires to be used at a later stage of this research, and adapting CL for College English teaching.

Piloting the Questionnaires

The two questionnaires of the LLOS and the SSS-CCEL (see the section on measures for details, pp. 100-108), were administered to the pilot class in the first session and re-administered after two weeks. The results indicated good test-retest reliability of above 0.8 for both of the questionnaires. The LLOS showed satisfactory internal consistency reliability for all subscales according to relevant Cronbach alpha values. As for the SSS-CCEL, an item (that is, “I often dominate the group and speak most of the time in group discussion”) was found to produce negative values for

the inter-item correlations with others. This signified that this item is not measuring the same underlying characteristics as other items (Pallant, 2007). In their responses to this item, the overwhelming majority of students expressed a high level of agreement that they did not dominate the group, which was substantially inconsistent with the responses to the other items on the subscale. The problem with this item may be due to the specific characteristics of Chinese students, who tend to show unwillingness to communicate in class and also have few opportunities to work in groups in traditional classrooms. In other words, the students' extremely positive self-report on this item did not necessarily mean they were equally participating in teamwork but might reflect that they were not actually well involved and engaged in teamwork. There was a possibility that the same answer was chosen by two very different students: one did not dominate the group because s/he knew the value of equal participation, while the other did not dominate the group for lack of adequate participation. Thus this item was clearly weak in measuring the latent variable. As a result, this item was replaced with a new item "I have a clear picture of my personal role in teamwork and participate actively", which turned out to have good inter-item correlations with others. In addition, students' feedback on the comprehensibility of items was collected after the re-administration, and accordingly some modifications were made on the Chinese-version questionnaires for an easier and more accurate understanding. The details of test-retest coefficients and internal consistency coefficients of the final version of the two questionnaires will be reported in the section on measures.

Piloting Cooperative Learning Techniques

The trial of CL with the pilot class started from the third week of the first semester when tests and re-tests of the questionnaires were completed. Baloché (1998) contends there are three basic principles for getting started with CL: start simple, start small and keep going. In the pilot study, CL was introduced gradually in terms of structural complexity, task difficulty and activity duration, in parallel with students' familiarity with CL and their improvement in cooperative skills. Simple structures were first used such as Listen (or Think)-Pair-Share which includes three steps: students

listening to a recording or thinking on a given topic individually, then taking turns to exchange ideas with their partners, and finally being selected to share their partners' ideas with the class. The activities based on simple structures were relatively easy to handle for beginners for three reasons. First, they mainly involved pair work requiring minimum social skills. Second, topics were specific and closely related to students' interests and current learning materials. Third, the length of activities was kept within ten minutes. When simple structures were successfully tried with students, longer activities based on more complex structures were gradually included.

As for the major components of the adaptation used in this research, Student-Team Achievement-Division (STAD) (Slavin, 1995) and the Structural Approach (Kagan, 1994) were selected for five reasons. First, when CL is used with Chinese College English learners, full attention should be given to a proper balance between CL and whole-class instruction, because students have long been exposed to rigid teacher-centred pedagogy in which peer interaction and learner autonomy are not traditionally advocated. STAD provides an ideal model for this situation because it involves both teacher presentation and teamwork (Slavin, 1995). Second, the Structural Approach to CL provides a wealth of structures, which are designed for different functions or domains of learning and can be applied to a wide range of teamwork of different complexity, themes and objectivities (Kagan, 1994). Third, Kagan's structures provide well-prescribed step-by-step techniques for teamwork, which are easy to follow and ideal for students inadequate in cooperative skills and independent learning experiences. In addition, these structures are also handy for novice teachers to start with. Fourth, both STAD and the Structural Approach are flexible enough for using in different curricular and teaching materials, and suit text-book based College English teaching. Fifth, combining STAD and the Structural Approach is likely to maximize the integration of CL basic principles and elements, which, according to Johnson et al. (1998), Kagan (1994) and Slavin (1995), include positive interdependence, individual accountability, promotive interaction, equal participant of group members, equal opportunities for success, development of social skills, and group processing. This will be further illustrated in the section on intervention procedure when the final version of the adaptation is elaborated.

Apart from focusing on STAD and the Structural Approach, the adaptation with the pilot group

also drew useful elements from four other major CL methods including Learning Together (Johnson et al., 1998), Jigsaw (Aronson & Patnoe, 1997), Complex Instruction (Cohen, 1994) and Group Investigation (Sharan & Sharan, 1994). For instance, in the process of forming groups, special attention was given to the use of formal groups which originates from Learning Together and stresses group rewards. This facilitated group cohesion among teammates during this short-term intervention of several months. In addition to principles of group formation, some elements of the five-step CL lesson design in Learning Together were also included in the adaptation, from specifying objectives for the lesson to setting up grading criteria and processing teamwork assessment. Moreover, some elements of Jigsaw, Complex Instruction and Group Investigation (GI) were adopted in designing team tasks. This can be reflected in the sample assignment of a team task (see appendix 3), which was composed of a number of questions, each of which having its particular focus on certain learning aspects or intellectual abilities (e.g. grammar, speaking, writing, reasoning or problem investigating skills, and the ability to work with computer and create visual aids for the team presentation). Using multiple-abilities team tasks and integrating investigation in teamwork are respectively basic features of Complex Instruction and GI. The assignment also incorporated task specification, a key element of Jigsaw, by assigning each team member a particular sub-task or question s/he was good at or interested in. At a later stage, teammates worked together to teach each other so that all knew the answers to all questions. The elements of Jigsaw were also used when teaching materials which could easily be divided into a number of parts (e.g. a reading text made up of a number of sections with each focusing on the brief introduction of a particular country). In this example, teams were assigned different sections to work on, aiming at becoming competent enough to teach the class the particular section they were assigned.

The pilot study provided opportunities for the researcher to observe and reflect on the actual use of CL, and collect feedback from students, so as to achieve a more suitable CL adaptation for College English teaching. In this process, four major problems that inhibited the effectiveness of CL were located and relevant solutions were worked out.

First, students were initially grouped by the teacher in order to achieve a high level of

intra-team heterogeneity and inter-team homogeneity in terms of their English proficiency, interest in English, major subjects, gender and home provinces. However, it was later found that familiarity and group cohesion were hard to build up among teammates from different departments because they had different class schedules and seldom met each other except when in the English classes. This also restricted the extent to which they could work together for team assignments after class. Therefore, students' preference for teammates and their after-class availability for team assignments were taken into account when grouping students.

The second problem was related to group size. Initially the researcher decided to use six-member groups rather than foursomes so as to reduce the number of groups and make it easier for the teacher to monitor groups in the large class. However, the classroom seating arrangement required students to move chairs around to form groups, which created much noise and also wasted precious class time. Moreover, individual accountability could be somewhat missing in big groups because it was hard to always involve six individual roles or sub-tasks in teamwork and time was often not enough for every member to speak in brief activities. Students' feedback also indicated that they preferred to work in smaller groups with familiar teammates. Thus foursomes were considered more suitable, and were used at the later stage of the pilot study apart from one five-member group .

Third, the teacher found that managing a class of 12 groups was no easy job, especially at the initial stage when many students lacked experience in cooperation and autonomous learning. In view of this, a captain was chosen by and for each team on condition that s/he was willing to help others and had good organizational skills. The role of the team captain turned out to be helpful in improving the quality and efficiency of teamwork since captains managed and led their own teams, like teacher assistants.

Fourth, according to the teacher's observation and students' feedback, some teammates, especially lower achievers, were not committed to team tasks, because team success was not formally recognized in the final evaluation of the course and high achievers tended to be counted on to complete their team tasks. As a solution to this problem, two countermeasures were taken. First, Numbered-Heads-Together (NHT), one of Kagan's structures, was adapted to consolidate

the integration of individual accountability and positive interdependence among teammates in this particular context. Second, it was considered essential that course evaluation should offer formal recognition of teamwork as well as rewards for teams' improvement and progress, so that all teammates could take team success seriously and felt motivated to work toward it. Details of these two measures will be elaborated in the section on intervention procedures.

Overall, this pilot study involved the process of trying out CL techniques, observing their actual use, collecting students' feedback, reflecting on problems, making modifications and conducting re-trials. Two fundamental principles were rigidly observed in the adaptation of CL methods. First, the adaptation had to incorporate key elements of CL because it is these elements that distinguish CL from other types of group work and lead to its effectiveness. Second, the particularity of the College English teaching situation had to be taken into account, for instance, large-class teaching, the strong influence of traditional teacher-centred methods, learners' lack of teamwork experience, and new curriculum requirements which focus on listening and speaking. There is no doubt that this pilot study provided a solid foundation for the creation of a suitable CL method for College English teaching and was indispensable to the next stage, the main study.

Main Study

The main study lasted for a semester of 18 weeks, from February to July 2008. Two level-one classes were randomly selected, one as the intervention group (N=52) and the other as the comparison group (N=48). A pre-test-post-test control group quasi-experimental design was employed. On the first day of class, consent forms were distributed for students to sign after a brief explanation of the experimental procedure, the purpose of the research and the CL techniques to be used. Students in both classes all agreed to participate in the research, so two questionnaires, the LLOS and the SSS-CCEL, were immediately administered as pre-tests. Then the teacher talked with students about a possible date to administer the CET as the pre-test. Because students were from 14 different majors with different class schedules, the weekend was the only time everybody could be available. Thus, on the first Saturday morning of the semester, the CET

written test was administered to both groups as the pre-test of listening, reading, writing and vocabulary. The spoken English test was conducted over the rest of the weekend, with six experienced College English teachers as examiners. The back-up CET paper for the previous final exam was used as the pre-test paper for this research. Teaching started from the second week: the intervention group was instructed using the adapted CL method, while the comparison group was taught using traditional instruction. Great care was taken to control extraneous variables (e.g. teaching content, learning resources and facilities, course assignments, teaching time arrangements, and the teacher's familiarity with the two different teaching methods), so that both groups were treated the same way in all respects except for the teaching method (see details in the following section on intervention procedure). At the end of the semester the LLOS, the SSS-CCEL and the CET were administered again to both groups as post-tests.

Intervention Procedure

A quasi-experimental process involves three key sequential stages: “establishing the baseline data; providing the intervention; and measuring the results” (Mutch, 2005, p. 117). It is obvious that the intervention plays a critical mediating role in linking the other two stages. In other words, it is only when a proper intervention is implemented that the differences between the results and the baseline data can be justified. Therefore, the intervention procedure forms a critical basis for the validity of an empirical study. In addition, an elaboration of the particular intervention employed in the study also enables others to conduct a true replication of the experiment and this forms an integral part of assessing the research reliability. However, in the CL research field, as Sharan (2002) states, it remains a critical problem that many researchers fail to make clear what particular methods have been employed, which inevitably puts a big question mark on the validity and reliability of their research results. The intervention procedure employed in this empirical study involved the use of the two different teaching methods: the adapted CL method used with the intervention group and traditional instruction used with the comparison group. Details of the two

methods will be elaborated in this section and contrasts between them will be specified.

The Adapted Cooperative Learning Method Used with the Intervention Group²

The adaptation was mainly made from Slavin's STAD and Kagan's Structural approach, but also included some useful elements from other methods, for instance, the incorporation of formal groups, multiple-abilities teamwork, task specification and investigation process, which were inspired respectively by Learning Together, Complex Instruction, Jigsaw and GI (details have been illustrated in the section on piloting CL techniques). The adapted CL method involved three aspects of the teaching: team formation, technique adaptation and course evaluation (see Figure 1).

Team Formation

Using CL in ELT entails appropriately teaming students of differing levels of language proficiency in a supportive environment so that promotive interaction can be generated. When forming teams, five factors must be taken into consideration: size, selection, composition, duration and organization (Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994).

Foursomes were used as the basis for teams for three reasons. First, the seating arrangement allowed pairs of students to turn around and form foursomes with two others behind them. Second, foursomes allowed pair work within a team, which doubled participation and lines of communication. Third, small teams were easy to manage for students, allowing for individual participation and accountability.

Three options exist for selecting team members: by teachers, by students themselves, or on a random basis. In this study, students were from different departments with different class schedules, and the English class was the only time they were all together. So teams were first selected by students according to their availability after class to meet for team assignments, as well as their preference for working together.

² This section on the adapted cooperative learning method has been published in *ELT Journal*, doi: 10.1093/elt/ccq021.

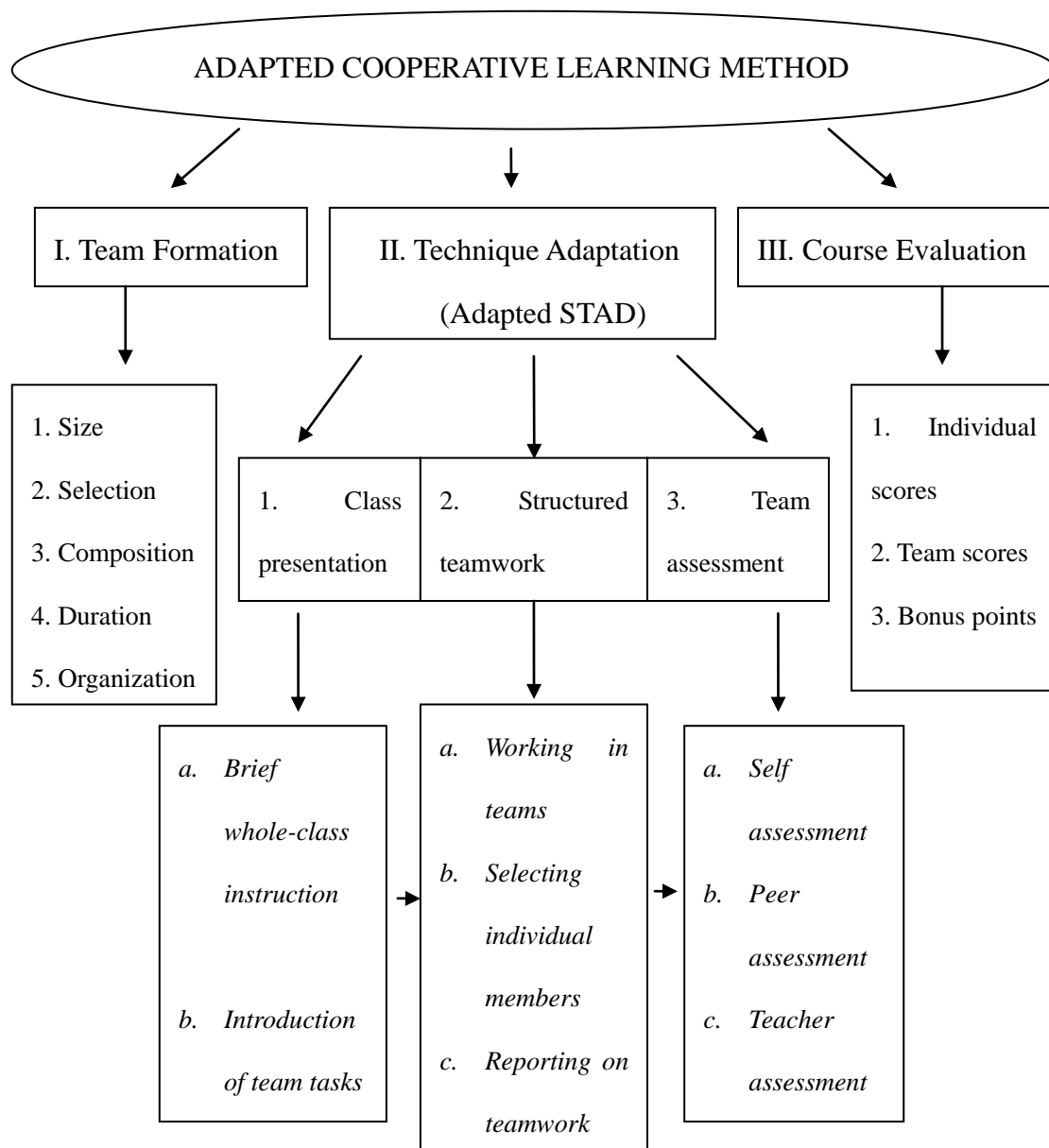


Figure 1: Components and procedures of an adapted CL method with Chinese EFL learners.

On the basis of student-selected teams, adjustments were made to achieve maximum heterogeneity especially in terms of language proficiency. Special efforts were made to build up inter-team comparability and academic balance by avoiding teams either composed of four high-achievers or four low-achievers, based on the results of the previous final exam. However, due to the restrictions of their class schedules, several teams were still somewhat homogeneous for lack of either a high-achiever or a low-achiever. In view of this, a remedial evaluation technique

emphasizing improvements in teams was applied to minimize the impact of inter-team differences on team success (see the section on course evaluation for details, p. 95).

The cooperative foursomes were used for the whole semester. This gave students who were unfamiliar with teamwork more time to develop cooperative skills, build group cohesion and overcome difficulties in working together. The use of long-term learning teams with stable membership is likely to enhance the quality and quantity of learning, improve class attendance, develop positive attitudes towards learning, and particularly suits the context of large class teaching where students have diverse abilities and needs (Johnson et al., 1998).

In addition, three steps were taken to organize teams to enhance team cooperation and cohesion:

- 1) Each team chose its own name, which was something all team members agreed on and could express their team identity.
- 2) The four members in each team were coded as Apple, Bean, Cat and Dog. (This way of coding was the students' choice while in most research, members are numbered off from one to four) Assigning each member a stable code was essential for implementing Kagan's (1994) NHT, which was modified and much used in this research.
- 3) Each team chose a team captain, who assisted the teacher by managing and leading his/her own team.

Technique Adaptation

A frequently used technique in this research was the combination of an adapted version of Slavin's (1995) STAD and a modification of Kagan's (1994) NHT. Since the frequent use of individual quizzes originally involved in STAD was impossible for such a big class of 52 students, STAD was adapted into a version involving three sequential components—class presentation, structured teamwork and team assessment. NHT was modified to organize and process teamwork, which will be explained in the section on structured teamwork.

1. Class Presentation

Class presentation by the teacher served as a basis for the structured teamwork and team assessment which followed. The presentation could be related to reading texts, writing skills, vocabulary, grammar, or a replay of audiovisual materials. The presentation took the form of whole-class teaching, but was brief compared with traditional teaching, because many learning materials were set aside to be completed by teamwork. For instance, a team task on figuring out where the narrator originally comes from (see appendix 4) derived from a reading text depicting her boyfriend's first meeting with her mum. The presentation on the text started with a five-minute whole-class brainstorming of possible responses of a Chinese mum to her daughter's plan to marry a foreigner. This provided a lead-in to engage students in the reading materials. Then the teacher spent 10 to 15 minutes focusing on the portions of the text involving unfamiliar background knowledge and difficult language points that might impede students' understanding. The remaining part which contained answers and clues to the narrator's gender (e.g. the way that the narrator introduced her boyfriend to her family, content of conversation between the narrator and her mum, and the way that her mum cooked the meal) was left for the teamwork, in which teammates put their heads together and read for detailed comprehension to complete the task.

Class presentation also included the introduction of cooperative team tasks undertaken immediately afterwards. A worksheet (see appendix 4) on the team task was usually distributed to each team. The worksheet typically included task requirements and some scaffolding phrases for team cooperation and task completion. Specific grading criteria were given as the rubric for students to follow when doing self/peer grading, which related to loudness, clarity, comprehensibility, and length of their speech, as well as adequate use of eye contact in speaking. Each team was also required to put on the worksheet its team name, individual tasks or roles, and agreed grades for presenting teams. The preparation of worksheets was considered essential for the successful use of CL activities at the initial stage. However, a simple team-grade sheet (see appendix 5) was used instead at times, for instance, when the team task was simple and brief (e.g.

a five-minute warm-up questions) and easy-to-follow, when students got accustomed to the commonly-used rubric or procedures for assessment, especially at the later stage, or when the team task and relevant requirements were included in the textbook (a few ready-to-use team tasks were included in the textbook for listening and speaking but rarely for reading and writing).

2. Structured Teamwork

The ultimate purpose of the teamwork design was to generate more peer interaction and meaningful negotiation in the process of completing designated tasks. More emphasis was placed on communicative fluency as the basis for linguistic accuracy. Students were encouraged to get meaning across instead of simply focusing on accuracy of language forms. NHT was modified to structure teamwork as follows:

- 1) Students put their heads together to work on tasks within given time limits. The task was based on textbook learning materials, and might be a five-minute class activity (e.g. brainstorming vegetable names) or a team assignment to be completed within a couple of weeks (e.g. preparing a ten-minute speech on “the brain drain in developing countries: reasons and results” based on their knowledge, information from textbooks and after-class cooperative research). A small amount of Chinese was allowed for team discussion but not for presenting work. The teacher was available to provide scaffolding when necessary.
- 2) A team was randomly selected as a presenting team, and then a particular code was randomly selected for the team.
- 3) The student with the code represented his/her team and reported on the team’s work in front of the whole class. The student’s performance was assessed and this grade was recorded as his/her team’s grade (This assessment process will be further elaborated in the next section).
- 4) More teams were chosen to report by repeating 2) and 3). Occasionally the teacher made the choice so that each team was given an equal opportunity for presentation and all teams had the same total number of assessments over the semester. Also, the appropriate use of the teacher’s choice could adjust the imbalance of “luck of the draw” and ensured that each teammate had

similar presenting chances over the semester.

NHT was modified mainly for two reasons. First, the original NHT expects every team to present, which does not suit large classes with limited teaching hours. Second, NHT originally chooses students with the same number in each team to report; consequently, while the first student is presenting, others with the same number are busy doing final preparation or rehearsal work instead of listening attentively. This impedes both classroom management and students' learning outcomes.

The purpose of this structure was having some teams present their work while not knowing in advance which teams would be selected, and having one student represent his/her team without knowing in advance who this person would be. This technique is aimed at facilitating involvement of all students when they individually hold the responsibility for team success. This is to ensure that students are highly motivated to learn and participate because they do not want to disadvantage their teammates due to their own inadequate work, and also to ensure that students are obliged to help each other learn because any teammate is potentially the team representative. This technique is aimed at integrating positive interdependence, individual accountability, promotive interaction, equal participation and cooperative skills in teamwork. It particularly suits large class teaching, where teachers find it difficult to monitor every student's performance. However, for the first couple of weeks in this study, students were allowed to volunteer answers in order to reduce anxiety and provide a model of what was expected.

3. Team Assessment

In contrast to traditional assessment where grades are simply decided by teachers, the four-step assessment process used in this study also embodied self-assessment and peer assessment:

1) Students worked in teams to assess presentations according to the criteria specified in advance. Audience teams, who did not have the chance to present, were required to put on their worksheets (or team-grade sheets) an agreed grade for each presenting team. Likewise, presenting teams graded themselves by reflecting on their own work. Grades ranged from C to A+.

2) The teacher provided her feedback on presentations by pointing out both strengths and weaknesses and giving some constructive suggestions on improving the work. (Peer/self-grading preceded the teacher's feedback so that it did not influence students' opinions.) At a later stage, students could volunteer to comment on peers' performance.

3) After all the selected teams had presented their work, worksheets or team-grade sheets were collected so that teams' self-grading and peer grading could be referred to by the teacher when making a decision on the grades of presenting teams. Collecting worksheets was also useful for the teacher to provide feedback on the audience teams' work from the notes recorded on their worksheets. This made audience teams feel their work was valued though not formally assessed and graded.

4) The teacher decided on the grades for presenting teams, which were usually calculated by averaging out the grades given by both the teacher and students. The final grades were notified by posting teamwork records (see appendix 6 for a sample sheet) on the classroom wall. If there was any disagreement from students, it would be discussed and the grade could be adjusted if necessary.

Both self- and peer assessment enable learners to reflect on their learning experiences and are integral parts of group processing. They enhance students' academic and social development, facilitate high-order thinking, and can create a favourable learning atmosphere of democracy and equality (Johnson et al., 1998). The requirement of peer grading is also helpful in keeping audience teams attentive and on-task during presentations so as to produce sound evaluations and comments afterwards. It was noted that students took the assessment very seriously and the results of peer grading, in most cases, turned out to be very close to the teacher's. In addition, this team assessment system is likely to achieve a good balance between formative and summative assessment through the teacher and peers providing immediate feedback and assigning team grades. This may largely remedy a weakness of the original STAD which focuses on summative assessment based on individual quizzes. Incorporating different forms of assessment in achievement evaluation can bring a wide variety of benefits to learning and teaching (for a review see the section on assessing cooperative learning group work, pp. 47-51).

Course Evaluation

Course evaluation recognized both team work and individual efforts. The students' final scores on the course were composed of two parts: 70 percent from individual scores on the final exam at the end of the semester, and 30 percent from team scores based on team grades. (It should be clarified here that the data used for evaluating students' language proficiency in the research reported in this thesis only included individual scores from the final examinations of CET, which ensured a high level of objectivity).

The team assessment system used in this study stressed equal opportunities for success by incorporating improvement points (Slavin 1995) as indicators of students' improvements. Slavin (1995) recommends using improvement points for individual students with their initial scores as the baseline scores to compare with. In this research, the technique of improvement points for individual students was adapted for teams by using the latest team grades as the baseline for adjusting the previous grade; that is, adjustments were made to previous team grades according to the extent of the improvement made at a particular time. For instance, team grades from previous work could be adjusted from B- to B, if they gained a B+ at a particular time (see appendix 6 for more examples). This adaptation enabled teams to compete with themselves rather than with others, and made this technique of improvement points more motivating due to a constantly updated team base grade. In other words, if a team got a C at a time, this would not be so depressing or devastating as in the traditional evaluation; on the contrary, the team knew they still had a chance to improve this low grade if they could get a good grade next time, so they would be motivated to try harder and improve their performance. Also, the use of improvement points helped to make up for possible inter-team gaps in overall academic level and facilitated equal opportunities for success between teams.

The teacher also introduced bonus points for teams into the assessment in order to encourage student participation and invite more voices into classroom teaching. Apart from the formal structured teamwork presented by selected teams, there were also many informal class activities (e.g. whole-class brainstorming on warm-up questions during class presentation) that needed

volunteer participation. The bonus point technique was specially designed for those cooperative teams that volunteered quick responses to the teacher's questions, offered comments on the performance of presenting teams, and shared ideas with classmates. Chinese students are usually afraid of being thought of as "show-offs" so they are likely to keep silent even if they have some good ideas. This bonus point technique gives teammates a good reason to encourage each other to speak out and practise English so as to win bonus points for their teams. Notably in this study, even some initially reticent students were prompted to volunteer and contribute to their team points under peer influence. This was highly conducive to building up a lively and comfortable atmosphere for both teaching and learning.

Traditional Instruction Used with the Comparison Group

The traditional techniques used with the comparison group mainly involved teacher-dominated whole-class instruction, which focused on the accuracy of vocabulary and grammar, and the processing of texts lexically and syntactically. Peer interaction, language practice and communicative fluency were not emphasized because teacher talk took up most of the class sessions. Discussion topics and learning tasks, which were carefully designed to suit students of different language levels within a cooperative team, were also modified for use with the comparison group mostly in the form of direct instruction or occasionally traditional group work.

For example, the CL team task cited in appendix 4 could be dealt with in the traditional classroom in either of two ways: direct instruction or traditional group work. In direct instruction the teacher would pose the question to students, then point out all the clues to the answer during the process of explaining the text in detail, and finally provide the answer herself. This way, there would be little peer interaction. In this particular case, the researcher actually chose to use traditional groups to process this task, which typically included three steps:

- 1) Students worked on the task with desk-mates or neighbours in groups which could involve students ranging from two to six.
- 2) Several students volunteered to present their answers.

- 3) The teacher provided feedback and comments on the students' presentation, and decided on grades for individual students who presented. These grades formed an indicator of the students' class performance.

Generally speaking, traditional group work differed from CL teamwork mainly in four aspects. First, traditional groups were formed by putting several nearest neighbours together on a random and temporary basis, while home teams with stable membership were used in the CL classroom. Second, traditional group work did not include two key elements of CL, that is, positive interdependence and individual accountability, which are essential for CL teamwork. Third, students in traditional groups worked together but volunteered to report on group work. Mostly it was a small number of high-achievers who took the opportunity to speak. Sometimes no students would volunteer, so the teacher had to provide the answer herself. This formed a contrast with cooperative group work where anyone stood a chance of being selected as a team representative to present the results of their teamwork. Fourth, in traditional group work students were assessed individually, with the teacher as the sole assessor, while students in CL teams were assessed as a team, based on the opinion of both the teacher and students. So in the course evaluation, although it was the same for both groups that 70% of the final scores on the course came from the final exam of the CET, the sources of the other 30% were different—it was individual grades on either class performance or assignments for the comparison group, while it was team grades for the intervention group.

Summary of Contrasts between Cooperative learning and Traditional Teaching

The contrasts between CL and traditional techniques used in this research can be summarised briefly as follows. First, whole-class direct instruction was brief for the intervention group in the CL classroom, while it took up most of the time in the traditional classroom with the comparison group. Second, CL activities incorporating positive interdependence and individual accountability were frequently used with the intervention group, while the comparison group was only occasionally exposed to traditional group work since direct instruction was mainly used with them.

Third, CL activities were aimed at increasing the production of peer interaction and meaningful negotiation among students, with emphasis on communicative fluency, while traditional teaching focused on accuracy of language forms through careful explanation of grammar, vocabulary, sentence structures and texts, as well as the use of a large number of repetitive drills. Fourth, in the CL classroom, the teacher acted as a facilitator, guiding, monitoring and observing students' efforts in learning, while students played an active role in teamwork and provided each other with comprehensible input and output. In the traditional teaching the teacher was class controller, language instructor and transmitter of knowledge, as well as the main provider of comprehensible input. Students mostly listened to the teacher and studied learning materials individually with little chance of meaningful communication with peers.

Contrasts between the two teaching methods can be further illustrated through an example of a typical 50-minute session, which gives a full picture of what was going on in the two different classrooms regarding teaching listening and speaking (see appendix 7). Additional examples are also provided to illustrate the contrasts regarding the teaching of writing (see appendix 8) and organizing vocabulary learning activities (see appendix 9). It should be noted that, due to the limited teaching hours, the writing process mostly occurred after class as an assignment, for which the intervention class worked in teams and students in the comparison class worked individually. In teaching vocabulary, what distinguished CL and traditional instruction typically involved three kinds of activities: extending lists, creating stories, and distinguishing confusing words or phrases (details can be seen in appendix 9). In addition, it was bonus points rather than team grades that were typically applied to vocabulary learning activities which were mostly very brief and encouraged volunteers' participation.

Measures

In line with the three research questions which were respectively related to the effects of the CL approach versus traditional instruction on students' language proficiency, learning motivation and

social skills, there were three instruments used in this research: the College English Test (CET) which was used to test five aspects of language proficiency, the Language Learning Orientations Scale (LLOS) measuring six areas of learning motivation, and the Social Skills Scale for Chinese College English Learners (SSS-CCEL) measuring eight areas of social skills.

College English Test

The CET (see appendix 10) was composed of two parts: a written test and a spoken test. It is a widely used means of assessing College English learning and teaching in China, either administered nationwide or within a university or college, and its reliability and validity have been well established (Yang & Weir, 1998). According to the information provided by the CET website (see http://www.cet.edu.cn/cet_teach2_3.htm#15), reliabilities measured by both KR20 and Cronbach's alpha are consistently above 0.85.

The CET is often administrated at the end of each semester as the final examination taken individually by students. The CET test item bank, established through years of effort by the Department of Foreign Languages at the University, is an important resource which teachers depend on to select test items. The design of the test paper is carefully planned, with rigorous procedures carried out to ensure content validity of items on the test. The design and composition of the test paper goes through three levels of examination and approval. Firstly, a team of experienced teachers are selected to decide on test items, each accountable for one part of the test, for example, listening or reading. Then they circulate the selected items within the team in order to agree on items for each part to compose a whole test paper. Secondly, the test paper is submitted to the course-coordinator to double-check that the items make up a good measure of learners' language proficiency. Finally, it goes to the head of the department for final checking and approval.

The design of the CET paper used in the present research followed the procedures outlined above. The written tests used in this research lasted 120 minutes, and comprised four aspects: listening, reading, vocabulary and writing, which respectively accounted for 25 percent, 30

percent, 20 percent and 15 percent of total scores (with the remaining 10 percent allocated to speaking). There were 25 items in the section on listening, 15 items on reading, and 40 items on vocabulary. Most of these items used a multiple-choice format, and the answer sheets were automatically scored by computer. For the other items, relating to spot dictation (a part of listening) and writing, two teachers of College English were assigned as examiners and scored separately in line with relevant criteria. For spot dictation, correct answers were provided for scoring. The scoring criteria for writing were related to its coherence and cohesion, grammatical range and accuracy, as well as the length of the writing. The two examiners reached agreement on a common score for each student through discussion.

The CET spoken test was conducted with small groups of four students arranged according to their order on the class register, and special attention was paid to avoiding the use of home teams in the test for the intervention class. The whole process of the speaking test took about 20 minutes for each group and mainly involved three parts conducted in English. In the first part, lasting about five minutes, the examiner asked students some questions about their background. Next, each student gave an individual presentation on a given topic for one and a half minutes, which was followed by a four-minute group discussion. Finally, in the last five minutes, the examiner asked more questions to further assess the oral communicative competence of each student. Six experienced College English teachers were assigned to test students' speaking ability. They worked in pairs, scoring separately against the criteria for the CET spoken test (see appendix 11), and then reaching consensus through discussion.

Language Learning Orientations Scale

The LLOS (see appendix 12) was developed by Dr. Kimberly Noels on the basis of self-determination theory. Self-determination theory postulates that “there are two general type of motivation, one based on intrinsic interest in the activity per se and the other based on rewards extrinsic to the activity itself” (Noels, Pelletier, Clement & Vallerand, 2000, p. 60). It also suggests that, in addition to intrinsic motivation and extrinsic motivation which are both intentional, there is

amotivation which stands in contrast and reflects a lack of intention and motivation (Gagne & Deci, 2005; Noels et al., 2000). Gagne and Deci (2005, p. 335) propose a self-determination continuum, which “ranges from amotivation, which is wholly lacking in self-determination, to intrinsic motivation, which is invariantly self-determined”. Along the continuum between amotivation and intrinsic motivation are four subtypes of extrinsic motivation: external motivation, introjected motivation, identified motivation and integrated motivation. External motivation derives totally from external pressures and is “characterized by performing an activity in order to achieve a reward or avoid a punishment” (Comanaru & Noels, 2009, p. 134). With introjected motivation, “one carries out an activity in order to temper internal pressures, particularly a sense of guilt, or for ego enhancement” (Comanaru & Noels, 2009, p. 134). Identified motivation refers to having greater volition to do an activity because people perceive the behaviour to be associated with their identities and reflect an aspect of themselves. With integrated motivation, people are fully aware that doing an activity is an integral part of who they are; in the specific context of language learning, integrated motivation is particularly linked to a desire to involve in the target community and identify with its members.

Following the framework described above, the LLOS consists of 42 items and is intended to measure six factors: intrinsic motivation, integrated motivation, identified motivation, introjected motivation, external motivation and amotivation. A seven-point Likert scale is used in the questionnaire, ranging from one to seven (respectively representing “does not correspond at all” to “corresponds exactly”). The LLOS is already available both in Chinese and English, and the researcher has Dr. Noels’ permission to use it as a measure in this research. According to students’ responses during the pilot study in 2007, some minor changes and adjustments have been made to the Chinese translation of several items for a better understanding, and the researcher has reported these changes to Dr. Noels.

The two sets of data respectively from the pre-test and the post-test suggested good internal consistency reliability for the LLOS. Cronbach’s alpha coefficients in the pre-test are 0.824 (total scale), 0.913 (intrinsic motivation), 0.910 (integrated motivation), 0.841 (identified motivation), 0.817 (introjected motivation), 0.817 (external motivation) and 0.846 (amotivation). In the

post-test, Cronbach's alpha coefficients are 0.855 (total scale), 0.930 (intrinsic motivation), 0.910 (integrated motivation), 0.839 (identified motivation), 0.814 (introjected motivation), 0.862 (external motivation) and 0.817 (amotivation).

Social Skills Scale for Chinese College English Learners

The SSS-CCEL (see appendix 13) was developed by the researcher under the guidance of her supervisors. The decision to develop a social skills scale particularly for this research was made after an extensive literature review both on a variety of social skills assessments and a wide array of social skills relevant to CL. This careful review did not locate any existing social skills measures that suited the purpose of this research. Meanwhile, this literature review also helped formulate a range of appropriate items which should be included in the measurement of social skills for this research.

Inappropriateness of Existing Social Skills Assessments for this Research

A detailed review of social skills assessments found that many existing instruments (e.g. the School Social Behaviour Scales, the Social Skills Rating System, the School Social Skills Rating Scale and the Social Behaviour Assessment Inventory) were related to children (for a review see Demaray & Ruffalo, 1995). Considering that the norm, focus, and complexity of social skills are distinctively different for children and adults, these instruments were not applicable to this research with tertiary students as subjects. There was also a Teenage Inventory of Social Skills, composed of two versions respectively for boys and girls (Inderbitzen & Foster, 1992). Since evaluating differences in social skills between genders was not the objective of this research, this inventory was not considered suitable either.

Four instruments were found to be related to college students and adults, which were the College Self-Expression Scale (Galassi, DeLo, Galassi & Bastien, 1974), the Adult Self-Expression Scale (Gay, Hollandsworth & Galassi, 1975), the Social Relations Survey (Lorr,

Youniss & Stefic, 1991) and the Social Skills Inventory (Riggio, 2003; Riggio & Carney, 2003). The former three focused particularly on the assessment of social assertiveness—an ability to express oneself openly, directly, and honestly. Assertiveness includes expressions of positive feelings and ideas such as love, affection, admiration and agreement, which are valued by CL. But on the other hand, some of the personal traits such as defensiveness, aggressiveness and dominance, which are favoured aspects in assertiveness, are not particularly valued by CL which emphasizes interpersonal support, empathy and equal participation. This was a major reason for considering them inappropriate for this research. In addition to this, the Adult Self-Expression Scale was designed for adults from 18 to 60 with a broad coverage of social situations and relations with strangers, family, friends, business partners and authority figures such as a boss or a superior, which were considered irrelevant to College English learners. Also the Social Relations Survey was made up of dichotomous questions, which, compared with rating scales, could build in respondent bias and fail to catch necessary subtlety and complexity of the responses (Cohen et al., 2007). The Social Skills Inventory was found to be a comprehensive one including six broad categories: emotional expressivity, emotional sensitivity, emotional control, social expressivity, social sensitivity and social control (Riggio & Carney, 2003). However, since it was considered that several months' introduction of CL methods into classroom ELT would not result in such a wide range of social outcomes and effects, this inventory was not considered suitable for this research either. Although none of the existing instruments were found to suit the purpose of this research, their theories, constructs and layouts all provided valuable insights which could be drawn upon to develop an appropriate measure of social skills. Developing an instrument specially related to the use of CL became a must for this research.

Social Skills Relevant to Cooperative Learning

Social skills are generally “related to assertiveness behaviour, social bonding, and interpersonal communication”, and involve not only specific behaviours but also some attitudes, perceptions and feelings which are closely linked to how one actually behaves (Lorr et al., 1991, p. 506).

However, “There is perhaps no adequate single definition of social skills. The variety and assortment of dimensions labelled as social skills is enormous” (Riggio, 1986, p. 649). This means the components of social skills may change with the specific research focus and objectivity.

In the field of CL, behaviours, perceptions and attitudes which facilitate team cooperation and supportive interrelationships are particularly emphasized and valued. A detailed literature review has found that social skills are often used interchangeably with communication skills, interpersonal skills, cooperative skills, collaborative skills, interactive skills, teamwork skills or group skills in many published works on CL (e.g. Baloché, 1998; Brown & Thomson, 2000; Hill & Eckert, 1995; Johnson et al., 1998; Kagan, 1994; Slavin, 1995). As for the components of social skills, some leading CL researchers have used a variety of categories. For instance, Kagan (1994) proposes a collection of 12 social roles in accordance with 12 particular social skills (e.g. encourager of others, gatekeeper of equalizing participation, coach for helping, checker for understanding, and cheerleader for team celebration). Johnson et al. (1998) categorize social skills into four levels: forming skills (i.e. basic skills to establish minimum norms for cooperation in teamwork), functioning skills (i.e. skills to build up positive interpersonal relations), formulating skills (i.e. task-related skills such as understanding learning materials and using reasoning strategies) and fermenting skills (i.e. higher level skills such as reconceptualising materials and handling controversies). Following the same criteria, Baloché (1998) and Hill and Eckert (1995) classify social skills in a very similar way apart from different wording of social skill levels (e.g. group forming skills are similar to getting-together skills; problem solving skills are similar to getting-it skills). Dishon and O’Leary (1998) and Brown and Thomson (2000) hold that the whole cohort of social skills relating to CL can be divided into two general types: task skills and working relationship skills (or alternatively termed as maintenance skills), the former focusing on skills needed to complete academic work while the latter emphasizing those needed to develop and sustain positive interrelationships. In addition, many researchers (e.g. Gillies & Ashman, 2003a; Johnson et al., 1998; Slavin, 1995) contend that important social outcomes of CL experience are also associated with students’ psychological health, which can be reflected by learners’ increased self-esteem, sense of belonging, peer acceptance and liking for others. The majority of

comparative studies between CL and traditional approaches have showed significant gains in psychological health favouring students in cooperative classrooms (Johnson et al., 1998; Kagan, 1994; Slavin, 1995).

Overall, these perspectives from CL researchers show that social skills in CL involve both specific behaviours (either relating to task completion or interrelationships) and perceptions, feelings or attitudes which direct the specific behaviours. Although CL researchers present a variety of different ways of categorizing social skills, a close study of the elaboration and breakdown of the social skills reveals that they overlap with one another to a great extent and the majority of components are quite similar in spite of different wording.

Development of the Social Skills Scale for Chinese College English Learners

The above detailed literature review generated a large collection of specific social skills relevant to CL, and questions focusing on these skills made up the initial pool of items for the SSS-CCEL. These items were then selected, modified, synthesized and re-categorized, in accordance with the characteristics of Chinese College English learners and possible effects of a short-term intervention of 18 weeks. Since College English learners are adults, some skills specifically relevant to children (e.g. role of team cheerleader) were excluded from this scale. Meanwhile, considering the limited power of a short-term classroom intervention, the SSS-CCEL focused on some specific aspects which could be affected by the intervention.

The first draft of the SSS-CCEL was composed of 56 items, which were categorized into eight subscales (which will be elaborated in detail in the next section). The questionnaire, both in an English and a Chinese version, was piloted in 2007 with six Chinese students who newly arrived in New Zealand to study undergraduate courses. Based on their responses and feedback, the scale was revised with a focus on three aspects. First, some items showing little discrimination were changed or deleted. Second, ambiguous or difficult wording was clarified for easy understanding. Third, a few possible double-questions were modified to follow the rule to ask only one question at a time (Cohen et al., 2007). This resulted in a second draft composed of 40 items, each of the

subscales compromised of five items. The Chinese version of the second draft was back-translated into English by a university senior lecturer who is fluent in both Chinese and English. As a result, some minor adjustments were made in wording for a better and easier comprehension. Then the revised Chinese version was piloted again by being emailed to seven tertiary students in China. This time the students completed the questionnaire within 15 minutes and no queries and problems were reported. In the pilot study in 2007, the questionnaire was further tested and re-tested with 49 college students in Mainland China within an interval of two weeks, which showed a good test-retest reliability coefficient of 0.88. However this pilot study located an item which showed a negative inter-item correlation, and was not adequate to capture potential differences among students; therefore this item was replaced (see the section on piloting the questionnaires for details, pp. 81-82).

A seven-point Likert scale, rather than the most commonly used five-point one, was used with the SSS-CCEL for two reasons. First, most people “would not wish to be called extremists” and “prefer to appear like each other in many respects” (Cohen et al., 2007, p. 327). This is especially true of Chinese who are traditionally influenced by the Golden Mean of Confucianism. Translated to the completion of rating scales, this means that College English learners are very likely to avoid the two extreme poles at either end, which may result in the reduction of two positions in the scale, and actually change five points into three points. In view of this, a seven-point scale was considered more appropriate to capture accurate data from students. Second, considering the SSS-CCEL and the LLOS (which is also a seven-point Likert scale) would be administered to participants at one time, the Likert points of these two measures should be consistent.

Constructs of the Social Skills Scale for Chinese College English Learners

The SSS-CCEL is made up of eight subscales of social skills: self-confidence, sense of cohesion, initiative in socialization, being positive, checking for understanding, equal participation and accountability, acceptance and empathy, and conflict management.

Self-confidence, is a major component of self-esteem and a very important indicator of

psychological health. It means one feels sure of his/her abilities and values, and believes s/he can achieve success through his/her own competence and efforts. This kind of belief is closely correlated with a wide range of factors such as self-concept, self-esteem, self-efficacy, self-acceptance, and self-worth (Johnson et al., 1998; Slavin, 1995). CL is believed to be able to enhance students' self-esteem, reduce their learning anxiety, generate a feeling of valuing themselves as competent and important individuals, and eventually promote psychological health (Johnson et al., 1998; Kagan, 1994; Slavin, 1995).

Sense of cohesion refers to how one feels being part of a group, which is another important factor regarding individual psychological health and is likely to result from positive interdependence. People with a strong sense of cohesion have a feeling of liking others and being liked by others and perceive their world as comprehensible. They are willing to take on difficult tasks, cope with stress and challenge, hold responsibility for group success, and offer peer support (Johnson & Johnson, 2003; Kagan 1994; Slavin, 1995).

Initiative in socialization, as a major component of assertiveness (Lorr et al., 1991) and social expressivity (Riggio & Carney, 2003), refers to one's ability and willingness to start communication and interaction with others. CL incorporates individual accountability, equal participation and proacademic peer norms into classroom teaching. It is therefore expected to stimulate students' to initiate interaction with each other (Gillies & Ashman, 2003b; Kagan, 1994), and promote constructive socialization (Johnson et al., 1998).

Being positive refers to specific behaviours among peers to offer encouragement, praise, and mutual support to fulfil group goals. It is the fundamental component of interrelationship or maintenance skills (Brown & Thomson, 2000; Dishon & O'Leary, 1998) and serves as an essential condition for team success. A good number of studies have indicated that incorporation of positive interdependence facilitates positive attitudes towards others and promotive interaction among peers (Gillies & Ashman, 2003a; Kagan, 1994; Slavin, 1995).

Checking for understanding is aimed at ensuring that all team members have a good understanding of each other and a good grasp of learning materials. It is a very important skill for successful interaction, fulfilment of group goals and academic gains (Johnson et al., 1998; Kagan,

1994). This skill involves both offering and asking for explanation, clarification, elaboration, illustration and summarization to make the point presented well understood. Checking for understanding is one of the most indispensable skills throughout the whole process of CL, from functioning groups, understanding materials, completing tasks and making assessments (Baloche, 1998; Brown & Thomson, 2000; Dishon & O'Leary, 1998; Johnson et al., 1998; Kagan, 1994).

Equal participation and individual accountability are two basic principles of CL, and are closely associated with each other; that is, individual accountability often results in equal participation among students and equal participation is essential to consolidate individual accountability. "Participation is an essential ingredient for student success; equal participation is an essential ingredient for success of all students", for which individual accountability provides a powerful means (Kagan, 1994, p. 4:10).

Acceptance and empathy refer to one's ability to accept and understand others with different personalities, academic levels, social backgrounds, and personal perspectives. Acceptance and empathy, as an integral part of social skills (Lorr et al., 1991; Riggo, 2003), play a crucial role in maintaining positive and supportive interrelations and establishing effective and productive teams (Brown & Thomson, 2000; Dishon & O'Leary, 1998), and are possible social outcomes from well-structured CL teamwork (Gillies, 2007; Gillies & Ashman, 2003a; Johnson et al., 1998; Slavin, 1995).

Conflict management involves skills concerning good listening, persuading, negotiating, reflecting, reconceptualising, higher-level reasoning, critical thinking and decision making, which are aimed at reaching a consensus or a constructive solution. Conflict management skills are regarded as important social skills in CL (Baloche, 1998; Hill & Eckert, 1995; Johnson et al., 1998; Kagan, 1994). "Conflict is frequent and probably inevitable within cooperative effort" and "the absence of conflict within a cooperative endeavour may indicate apathy toward the task and each other" (Johnson & Johnson, 1994b, p. 66). Actually, conflicts may generate higher-quality teamwork and more individual academic gains if they are well managed.

To summarize, the SSS-CCEL, a seven-point Likert scale, was designed to assess differences an 18-week CL intervention could possibly bring to Chinese College English learners' social skills.

It consists of 40 items categorized into eight subscales, with five items comprising each subscale. When administering the questionnaire, the 40 items are rearranged so that every eighth item belongs to the same scale. For example, items 1, 9, 17, 25 and 33 assess self-confidence. The seven-point scale is coded from one to seven (respectively representing “not at all like me” to “exactly like me”) for scoring. 11 items require reverse scoring. Generally speaking, the higher students’ scores are, the better social skills they have.

The two sets of data respectively from the pre-test and the post-test showed satisfactory internal consistency reliability, according to the criteria that an ideal Cronbach alpha coefficient of a scale should be above 0.7 (DeVellis, 2003). Meanwhile, Pallant (2007, p. 95) states that “it is common to find quite low Cronbach values (e.g. .5)” for short scales with fewer than ten items. In the pre-test of this study, the Cronbach’s alpha coefficients are 0.855 (total scale), 0.624 (self-confidence), 0.635 (sense of cohesion), 0.694 (initiative in socialization), 0.726 (being positive), 0.677 (checking for understanding), 0.617 (equal participation and accountability), 0.714 (acceptance and empathy), and 0.738 (conflict management). In the post-test, the Cronbach’s alpha coefficients are 0.905 (total scale), 0.707 (self-confidence), 0.816 (sense of cohesion), 0.695 (initiative in socialization), 0.800 (being positive), 0.768 (checking for understanding), 0.801 (equal participation and accountability), 0.661 (acceptance and empathy), and 0.760 (conflict management). However, since the SSS-CCEL has only been used in this present study, it needs to be further tested for both reliability and validity before others could use it.

Data Analysis

The data collected for analysis in this research comprised pre-test and post-test scores on the CET, the LLOS and the SSS-CCEL. Specifically speaking, there were a total of 19 testing areas covered by three measures: five areas of language proficiency on the CET (i.e. listening, speaking, reading, writing, vocabulary), six areas of learning motivations on the LLOS (intrinsic motivation,

integrated motivation, identified motivation, introjected motivation, external motivation and amotivation), and eight areas of social skills on the SSS-CCEL (self-confidence, sense of cohesion, initiative in socialization, being positive, checking for understanding, equal participation and accountability, acceptance and empathy, and conflict management).

The 15th version of SPSS was used to conduct data analysis on the effects of the adapted CL method, versus traditional instruction, on each of the areas. According to Garson (2008, 2009) and Pallant (2007), there are three common ways of handling such data based on a pre-test-post-test control group quasi-experimental design. The first is one-way ANOVA on post-test scores. However, this involves ignoring the initial differences in pre-test scores, and is therefore not recommended. The second is mixed between-within subjects ANOVA, which is alternatively called repeated measures ANOVA. This ANOVA model has time as the within-subjects factor, pre-test and post-test scores as within-subjects variables, and type of group as the between-subjects variable. It can reveal how the groups have changed from pre-test to post-test according to the extent of interaction between time and group. The third is one-way ANCOVA, which is preferred when there are pre-test differences between groups. This is because this ANCOVA model has pre-test scores as the covariate, post-test scores as the dependent variable, and type of group as the fixed factor. In other words, this model holds constant any differences in the pre-test scores and evaluates the post-test differences between groups.

To detect the pre-test differences between the intervention and the comparison group, an independent-samples t-test was run to compare the two groups as regards their pre-test scores in each of the 19 areas on the three measures, as well as their ages and years of learning English. No significant differences were found between the two groups at pre-test since alpha values were greater than 0.05 for all the comparisons (see appendix 14). This finding indicated that the two groups were initially at a similar level in terms of their ages, experience of learning English, language proficiency, learning motivation and social skills. However, it was also noted that alpha values in several areas were quite close to the cut-off point, for instance, 0.060 in speaking and 0.054 in acceptance and empathy. Therefore, it was ideal for this research to use both mixed between-within subjects ANOVA and one-way ANCOVA, if the data met the assumptions required

for use of both statistical techniques.

Preliminary assumption testing was conducted in each area for the assumptions of normality, linearity, as well as homogeneity of variances, of inter-correlation and of regression slopes concerning the covariate and the dependent variable. It was found that assumptions were generally met except for violations of homogeneity of variances in such areas as sense of cohesion, being positive, acceptance and empathy, as well as integrated motivation. Fortunately, failure to meet the assumption of homogeneity of variances is not fatal to both mixed between-within subjects ANOVA and one way ANCOVA, and they are relatively robust, particularly when groups are of similar sample size (Garson, 2009; Norman, 2010). In this research, the numbers of participants in the two groups are of very similar size: one was 52 and the other 48. This confirmed that both mixed between-within subjects ANOVA and one-way ANCOVA could be used to analyze the data for this research.

Considering that the t-test is also a robust procedure against violations of equal variances (Larson-Hall, 2010; Pallant, 2007), paired-samples t-tests were conducted to evaluate differences between pre-test and post-test scores on the three measures for both groups in each of the areas. Effect size statistics were calculated for each group to gain an indication of the magnitude of differences between pre-test and post-test scores. Cohen's *d* which is the most commonly used effect size statistic for comparing two groups (Cohen, 1988; Larson-Hall, 2010) was used in this analysis. Since the *d* value is not directly provided by the SPSS, Pallant (2007) recommends a very useful website that provides a quick and easy way to calculate Cohen's *d* at <http://www.uccs.edu/~faculty/lbecker/>. In this research, the effect size statistics were calculated through entering mean scores and standard deviations of both pre- and post-test into this online calculator. Guidelines for interpreting the *d* value vary according to different academic fields and research purposes. According to Cohen (1998), effect sizes of around 0.2 are small, 0.5 moderate, and 0.8 large. Slavin (1990, 1995) considered effect sizes over 0.25 to be educationally meaningful. Recently, based on a synthesis of over 800 meta-analyses relating to achievement in the field of education, Hattie (2009) found 0.4 to be the average effect size for educational interventions, suggesting that an effect size above 0.4 is above average and therefore worth noting

when evaluating educational outcomes. Since Hattie's guideline is based on the latest research and is generated from a great number of educational studies, it is considered the most appropriate one to use for this research. Additionally, mean plots for each area on the three measures were also produced, allowing a visual inspection of the changes in mean scores of the two groups from pre-test to post-test.

To sum up, data from this research were analyzed by means of a range of statistical techniques, which include paired-samples t-tests, effect sizes, mixed between-within subjects ANOVAs, one-way ANCOVAs and mean plots. It is considered that the combined use of a variety of analysis techniques is conducive to providing richer information about the research, and enables the researcher to have a clearer picture about the phenomenon of interest. Therefore, it contributes to obtaining more accurate and reliable results, and reduces the possibility of biased findings due to the use of a single form of statistical analysis.

CHAPTER FOUR: RESULTS

Overview

This chapter presents the results based on data analyses of the three measures: the College English Test (CET), the Language Learning Orientations Scale (LLOS), and the Social Skills Scale for Chinese College English Learners (SSS-CCEL). In general, the results focus on seven aspects: mean scores, standard deviations of pre- and post-tests for each group, effect sizes using Cohen's d from the pre- to post-test for each group, alpha values of paired-samples t-tests for each group, alpha values of interaction effects between group and time from mixed between-within subjects ANOVAs, alpha values of post-test differences between groups from one way ANCOVAs, and mean plots for each of testing areas on the three measures.

Results from the College English Test³

Results obtained from the analysis of scores on the CET are summarized in Table 2, which includes mean scores, standard deviations, effect size statistics, and p values from paired-samples t-tests within each group, the ANOVAs and the ANCOVAs, in each of the five areas and the total score on the CET.

³ The effectiveness of cooperative learning in developing students' English language proficiency, based on the results from the College English Test, has been published in *Effective Education*, 2(2), 99-116.

Table 2: Summary of results on the CET for the intervention and comparison groups

CET	Group	Pre/ Post	Mean	S.D.	Sig. t-test	Effect size	Sig. ANOVA	Sig. ANCOVA
Listening	Int	Pre	14.750	3.458	0.000	0.747	0.037	0.013
		Post	17.231	3.178				
	Com	Pre	14.896	4.274	0.031	0.295		
		Post	16.000	3.135				
Speaking	Int	Pre	6.906	1.151	0.000	0.739	0.000	0.000
		Post	7.721	1.053				
	Com	Pre	6.527	0.790	0.080	0.160		
		Post	6.654	0.795				
Reading	Int	Pre	21.577	3.826	0.003	0.446	0.086	0.047
		Post	23.192	3.390				
	Com	Pre	21.708	4.267	0.802	0.040		
		Post	21.875	4.113				
Writing	Int	Pre	8.346	2.132	0.000	0.835	0.560	0.267
		Post	10.154	2.200				
	Com	Pre	8.000	2.415	0.000	0.713		
		Post	9.583	2.009				
Vocabulary	Int	Pre	14.106	2.042	0.000	0.809	0.606	0.551
		Post	15.596	1.615				
	Com	Pre	13.344	2.632	0.000	0.689		
		Post	15.063	2.351				
Total	Int	Pre	65.685	8.948	0.000	0.946	0.009	0.001
		Post	73.894	8.407				
	Com	Pre	64.475	10.446	0.000	0.485		
		Post	69.175	8.854				

In the area of listening, results of paired-samples t-tests showed both the intervention group [$t(51) = 5.813, p < 0.0005$] and the comparison group [$t(47) = 2.227, p = 0.031$] improved significantly from pre-test to post-test. The effect size for the intervention group ($d = 0.747$), according to Hattie's (2009) guidelines, was above average, but this was not the case for the comparison group ($d = 0.295$). Results of the ANOVA indicated a significant interaction effect for time and group [$F(1, 98) = 4.468, p = 0.037$] which suggests that the intervention group improved significantly more than the comparison group in listening. Likewise, results of the ANCOVA [$F(1, 97) = 6.401, p = 0.013$] showed a significant difference in post-test scores between the two groups after adjusting for pre-test scores. Further evidence for the difference between the two groups in listening competence is supported by inspection of the mean plots for listening (see Figure 2) which show that while both groups improved, the intervention group clearly improved substantially more than the comparison group.

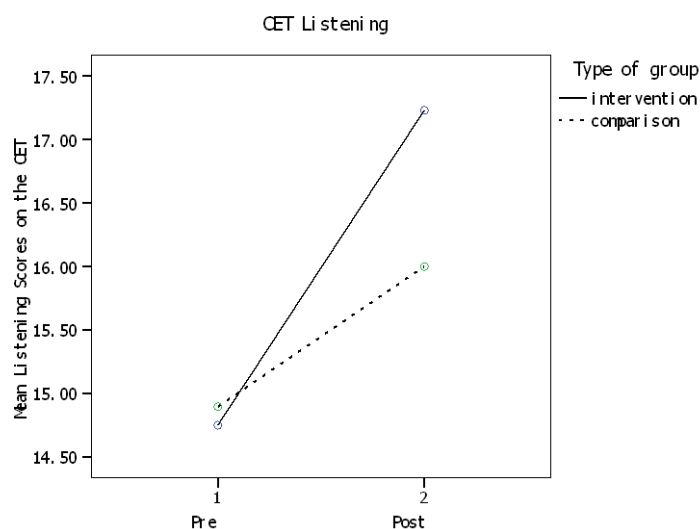


Figure 2: Pre- and post-test mean listening scores on the CET

In the area of speaking, the intervention group [$t(51) = 15.366, p < 0.0005$] improved significantly from pre-test to post-test, but the comparison group [$t(47) = 1.790, p = 0.080$] did not. The effect size for the intervention group ($d = 0.739$) was above average, but not for the comparison group ($d = 0.160$). Moreover, there was a significant interaction effect for time and group [$F(1, 98) = 61.502, p < 0.0005$] which indicates that the intervention group improved significantly more than the comparison group in speaking. Results of the ANCOVA [$F(1, 97) = 78.363, p < 0.0005$] told the same story that the intervention group improved significantly more than the comparison group on the post-test with differences in the pre-test held constant. The mean plots for speaking (see Figure 3) also show a substantial difference between the two groups in the degree of their improvements.

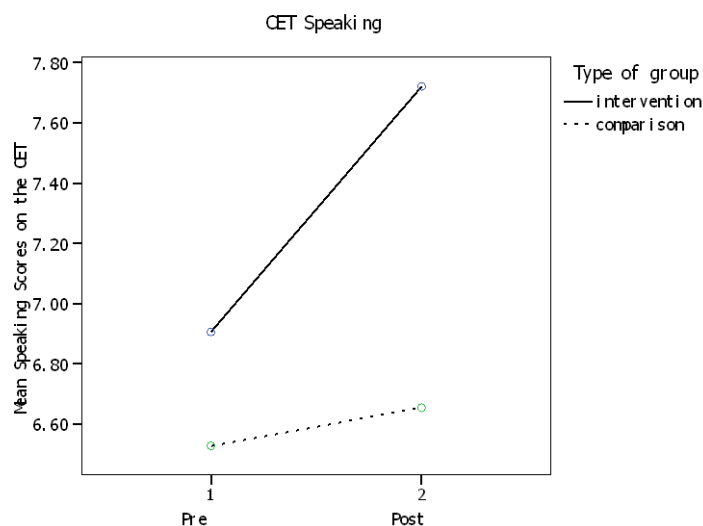


Figure 3: Pre- and post-test mean speaking scores on the CET

In the area of reading, results of the t-test analysis showed that significant improvement occurred in the intervention group [$t(51) = 3.100, p = 0.003$] but not in the comparison group [$t(47) = 0.253, p = 0.802$]. The effect size for the intervention group ($d = 0.446$) was above average but was extremely small for the comparison group ($d = 0.040$). Although the interaction effect [$F(1, 98) = 3.015, p = 0.086$] was not quite statistically significant, it suggests a trend of greater gains in the intervention group. Interestingly, results of the ANCOVA [$F(1, 97) = 4.036, p = 0.047$] showed significant post-test differences between groups with pre-test scores as the covariate. However, considering that this alpha value is very marginal relating to a cut-off point of 0.050, some caution must be exercised when making a conclusion. In this case, it is better to cautiously state that the intervention group substantially improved their reading competence more than the comparison group, but probably not at a statistically significant level yet. Further evidence for this situation can be observed in the mean plots for reading (see Figure 4), which show that the intervention group improved considerably while the comparison group remained almost at the same level.

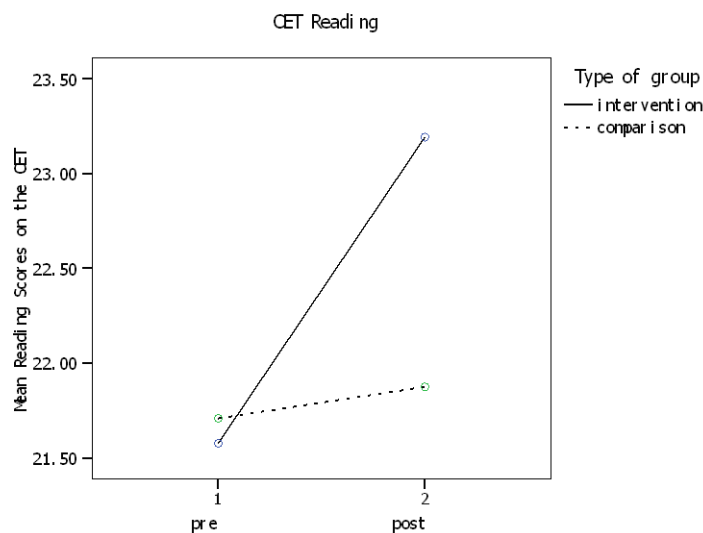


Figure 4: Pre- and post-test mean reading scores on the CET

In the area of writing, both the intervention group [$t(51) = 6.861, p < 0.0005$] and the comparison group [$t(47) = 5.674, p < 0.0005$] improved significantly. The effect size for the intervention group ($d = 0.835$) and the comparison group ($d = 0.713$) were both above average. The ANOVA showed no significant interaction effect [$F(1, 98) = 0.342, p = 0.560$] which indicates that there was no significant difference in improvement over time between the two groups in writing. Results of ANCOVA [$F(1, 97) = 1.234, p = 0.267$] indicated a similar situation in that the two groups exhibited no significant difference in their writing scores on the post-test, after adjusting for pre-test scores. Further evidence for the similarity in improvement of the two groups is indicated the mean plots for writing, which are composed of two rising lines that are nearly parallel (see Figure 5).

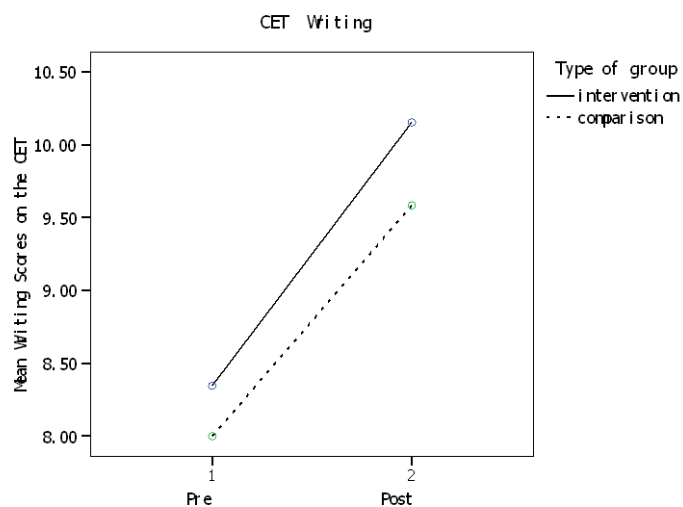


Figure 5: Pre- and post-test mean writing scores on the CET

In the area of vocabulary, results of the t-tests showed significant improvements for both the intervention group [$t(51) = 5.757, p < 0.0005$] and the comparison group [$t(47) = 4.714, p < 0.0005$]. Effect sizes for the intervention group ($d = 0.809$) and the comparison group ($d = 0.689$) were both above average. The ANOVA found no significant interaction effect [$F(1, 98) = 0.267, p = 0.606$] which indicates the two groups showed no significant difference in terms of improvements in vocabulary. The ANCOVA showed very similar results [$F(1, 97) = 0.359, p = 0.551$], which means there was no significant difference in the post-test scores between the two groups. This is further supported by inspection of the mean plots for vocabulary, which display two approximately parallel lines (see Figure 6).

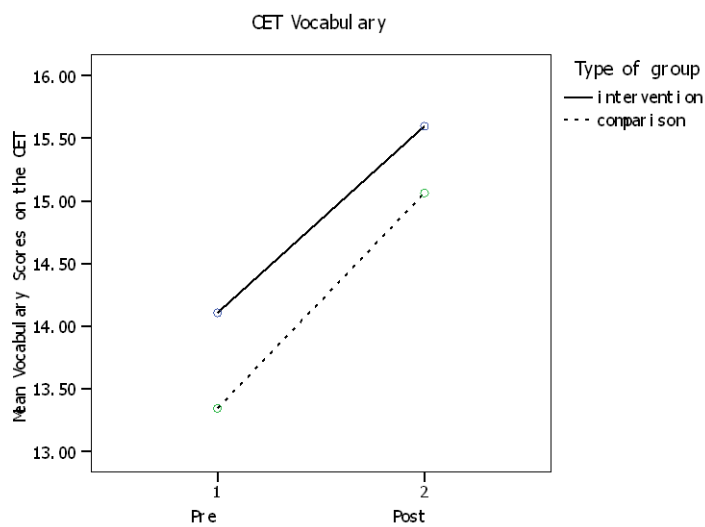


Figure 6: Pre- and post-test mean vocabulary scores on the CET

As for the CET total, results of paired-samples t-tests showed that both the intervention group [$t(51) = 10.271, p < 0.0005$] and the comparison group [$t(47) = 4.399, p < 0.0005$] improved significantly from pre-test to post-test. Effect size for the intervention group ($d = 0.946$) nearly doubled that for the comparison group ($d = 0.485$), although they were both above average. According to the ANOVA, there was a significant interaction effect for time and group [$F(1, 98) = 7.057, p = 0.009$] which suggests that the intervention group improved significantly more than the comparison group in overall English proficiency. Consistent with the ANOVA, results of the ANCOVA [$F(1, 97) = 11.456, p = 0.001$] showed a significant difference on the post-test between the two groups with the pre-test scores as a covariate control. Therefore, it can be safely stated that while both groups significantly improved their English language proficiency from pre-test to post-test, the intervention group made significantly more improvements than the comparison group. Further evidence for the difference between the two groups in total scores on the CET is provided by the mean plots (see Figure 7), which clearly show the different gradients of the two lines.

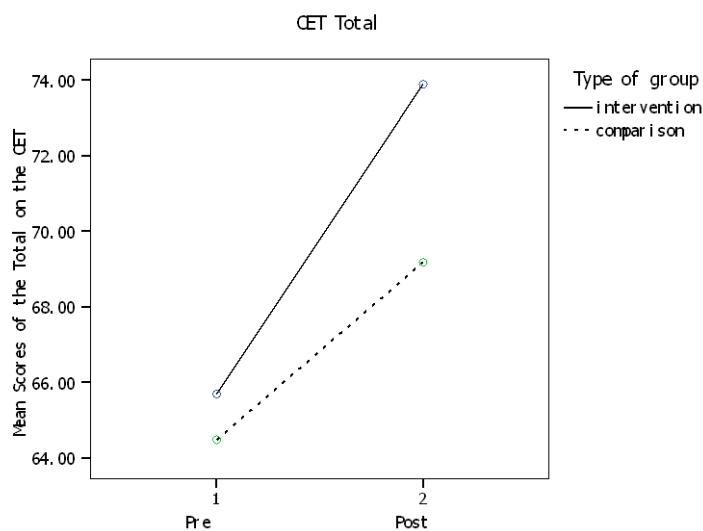


Figure 7: Pre- and post-test mean scores of the total on the CET

To sum up, the findings indicated that both groups made significant improvements between pre and post testing in their overall English language proficiency. However, when focusing on specific areas of language competence, the intervention group showed a significant increase in all areas whereas the comparison group did not increase much in speaking and reading. More importantly, it was found that the improvements made by the intervention group were significantly greater than those made by the comparison group in total scores on the CET as well as in the areas of listening and speaking, while not much different in writing and vocabulary. As for the area of reading, the between-group difference in terms of their improvements was of marginally statistical significance. Effect size statistics provided further evidence of specific effects of CL in teaching listening, speaking and reading, since effect sizes in these three areas were all above average for the intervention group but not for the comparison group.

Results from the Language Learning Orientations Scale

Results obtained from the analysis of scores on the LLOS are summarized in Table 3, which includes mean scores, standard deviations, Cohen's *d*, and *p* values of paired-samples *t*-tests within each group, the ANOVAs and the ANCOVAs, in each of the six areas and the total scores on the LLOS.

Table 3: Summary of results on the LLOS for the intervention and comparison groups

LLOS	Group	Pre/ Post	Mean	S.D.	Sig. t-test	Effect size	Sig. ANOVA	Sig. ANCOVA
Intrinsic motivation	Int	Pre	36.539	9.022	0.000	0.469	0.061	0.053
		Post	41.269	11.058				
	Com	Pre	36.563	11.616	0.067	0.164		
		Post	38.417	10.983				
Integrated motivation	Int	Pre	28.846	5.539	0.060	0.200	0.853	0.875
		Post	30.135	7.263				
	Com	Pre	28.667	9.762	0.064	0.164		
		Post	30.164	8.495				
Identified motivation	Int	Pre	27.539	5.717	0.038	0.183	0.797	0.711
		Post	28.577	5.651				
	Com	Pre	27.354	6.887	0.279	0.128		
		Post	28.167	5.751				
Introjected motivation	Int	Pre	22.154	6.102	0.093	0.167	0.979	0.943
		Post	23.192	6.331				
	Com	Pre	22.292	7.074	0.123	0.157		
		Post	23.354	6.446				
External motivation	Int	Pre	33.154	6.044	0.462	0.094	0.310	0.145
		Post	33.769	6.955				
	Com	Pre	31.604	7.249	0.485	-0.078		
		Post	31.021	7.617				
Amotivation	Int	Pre	23.039	4.121	0.111	0.165	0.622	0.654
		Post	23.615	2.716				
	Com	Pre	22.833	4.459	0.135	0.224		
		Post	23.750	3.693				
Total	Int	Pre	171.270	24.983	0.001	0.318	0.369	0.320
		Post	180.558	32.935				
	Com	Pre	169.313	37.631	0.082	0.154		
		Post	174.854	34.054				

In the area of intrinsic motivation, results of paired-samples t-tests showed that the intervention group [$t(51) = 4.168, p < 0.0005$] increased significantly from pre-test to post-test, but the comparison group [$t(47) = 1.878, p = 0.067$] did not. According to Hattie's (2009) guidelines, the effect size for the intervention group ($d = 0.469$) was above average, but not for the comparison group ($d = 0.164$). The ANOVA showed an interaction effect for time and group [$F(1, 98) = 3.605, p = 0.061$], which was close to a statistically significant level. Likewise, the ANCOVA generated a very marginal alpha value [$F(1, 97) = 3.845, p = 0.053$]. This shows a clear trend of greater increases in intrinsic motivation for the intervention group. However, the difference between the two groups in improving their intrinsic motivation did not reach a statistically significant level yet. The mean plots for intrinsic motivation (see Figure 8) also show a substantial difference in the magnitude of improvements between the two groups—they started at a very similar level but a substantial difference was evident by the end.

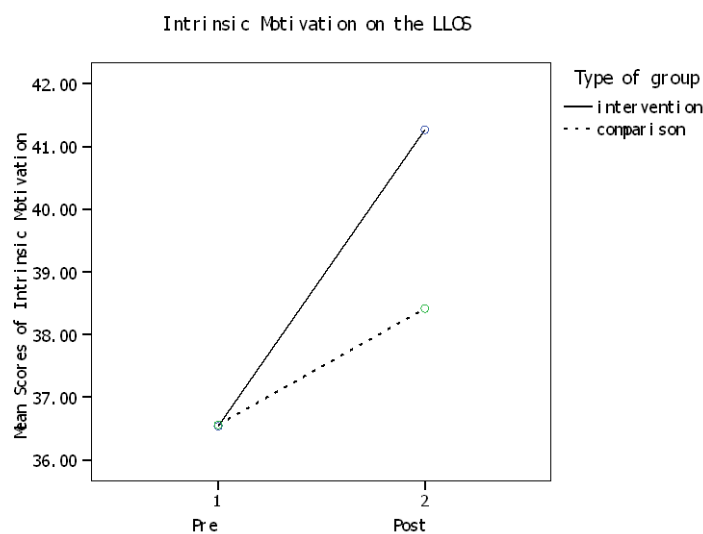


Figure 8: Pre- and post-test mean scores of intrinsic motivation on the LLOS

In the area of integrated motivation, both the intervention group [$t(51) = 1.921, p = 0.060$] and the comparison group [$t(47) = 1.897, p = 0.064$] improved from pre-test to post-test, but neither of them reached statistical significance. Effect sizes both for the intervention group ($d = 0.200$) and for the comparison group ($d = 0.164$) were below average. Results of both the ANOVA [$F(1, 98) = 0.035, p = 0.853$] and the ANCOVA [$F(1, 97) = 0.025, p = 0.875$] indicated no significant difference between the two groups in terms of their improvements in the area of integrated motivation from pre-test to post-test. The mean plots for integrated motivation (see Figure 9) also indicate that the magnitude of improvements between the two groups did not differ much.

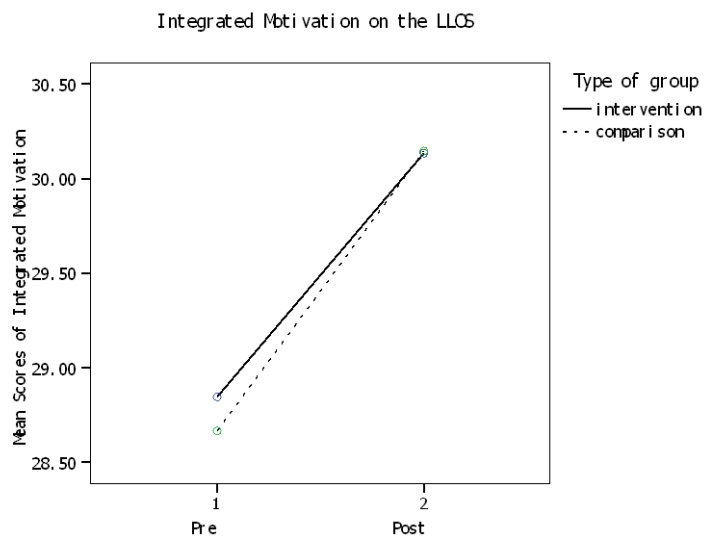


Figure 9: Pre- and post-test mean scores of integrated motivation on the LLOS

In the area of identified motivation, the intervention group [$t(51) = 2.127, p = 0.038$] improved significantly from pre-test to post-test, but the comparison group [$t(47) = 1.096, p = 0.279$] did not. Effect sizes for the intervention group ($d = 0.183$) and the comparison group ($d = 0.128$) were both small and below average. The ANOVA [$F(1, 98) = 0.067, p = 0.797$] and the ANCOVA [$F(1, 97) = 0.138, p = 0.711$] both indicated that there was little difference between the two groups in terms of their improvements from pre-test to post-test, although the results of t-tests were in favour of the intervention group. The mean plots for identified motivation (see Figure 10) also show that the magnitude of improvements between the two groups did not differ much.

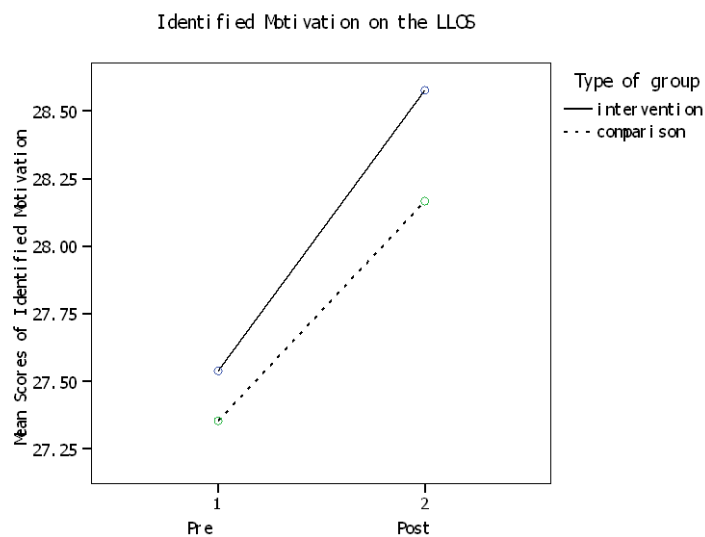


Figure 10: Pre- and post-test mean scores of identified motivation on the LLOS

In the area of introjected motivation, the results of paired-samples t-tests indicated no significant difference between the two time periods for both the intervention group [$t(51) = 1.712$, $p = 0.093$] and the comparison group [$t(47) = 1.569$, $p = 0.123$]. Effect sizes for both groups—the intervention group ($d = 0.167$) and the comparison group ($d = 0.157$)—were very similar and well below average. This situation was once again supported by results of the ANOVA [$F(1, 98) = 0.001$, $p = 0.979$] and the ANCOVA [$F(1, 97) = 0.005$, $p = 0.943$], which mean there were little between-group differences regarding their improvements in the area of introjected motivation. The mean plots for introjected motivation (see Figure 11) demonstrate this similarity by presenting two almost parallel lines.

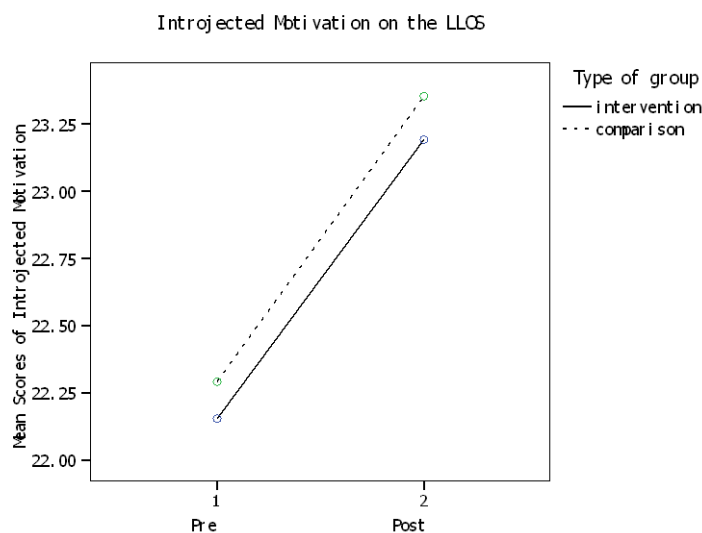


Figure 11: Pre- and post-test mean scores of introjected motivation on the LLOS

In the area of external motivation, it was noted that the intervention group showed slight gains from pre- to post-test while the comparison group experienced a slight decline. However, according to paired-samples t-tests, there was no significant difference between pre and post testing for both the intervention group [$t(51) = 0.742, p = 0.462$] and the comparison group [$t(47) = -0.704, p = 0.485$]. This is supported by the effect size statistics for the intervention group ($d = 0.094$) and the comparison group ($d = -0.078$), which were both well below average. Results of both the ANOVA [$F(1, 98) = 1.042, p = 0.310$] and the ANCOVA [$F(1, 97) = 2.155, p = 0.145$] indicated that there was no significant difference between the two groups as regards changes in external motivation from pre- to post-test. The mean plots for external motivation (see Figure 12) provide further support for the above findings by showing two lines wide apart without sign of interaction.

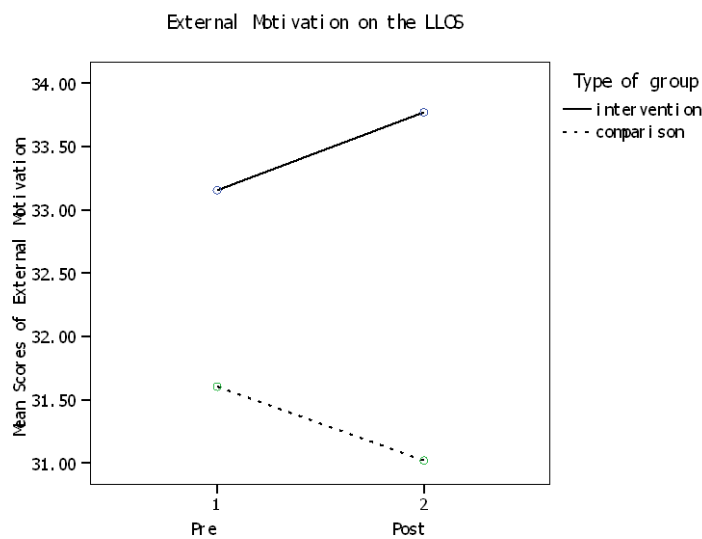


Figure 12: Pre- and post-test mean scores of external motivation on the LLOS

In the area of amotivation, it was found that the comparison group performed slightly better than the intervention group from pre- to post-test. But results of the t-tests showed no significant improvements for both the intervention group [$t(51) = 1.624, p = 0.111$] and the comparison group [$t(47) = 1.523, p = 0.135$]. Moreover, effect sizes for the intervention group ($d = 0.165$) and the comparison group ($d = 0.224$) were both below average. The similarity in this area between the two groups was further confirmed by results of both ANOVA [$F(1, 98) = 0.245, p = 0.622$] and ANCOVA [$F(1, 97) = 0.202, p = 0.654$]; that is, the two groups did not differ much in their improvements in this area between pre and post testing. This is further supported by inspection of the mean plots for amotivation (see Figure 13), which display a very slight interaction between the two lines.

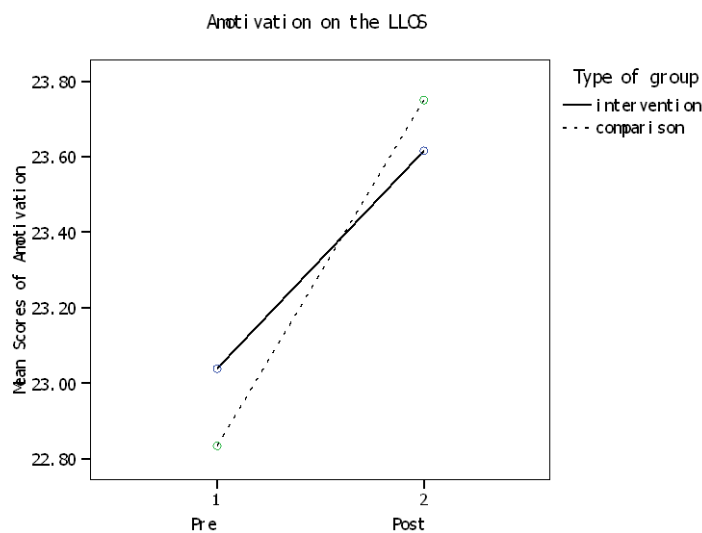


Figure 13: Pre- and post-test mean scores of amotivation on the LLOS

As for total scores on the LLOS, results of paired-samples t-tests showed that the intervention group [$t(51) = 3.370, p = 0.001$] significantly improved from pre- to post-test while the comparison group [$t(47) = 1.778, p = 0.082$] did not. The effect size for the intervention group ($d = 0.318$) approached an average level, but this was not the case for the comparison group ($d = 0.154$). Results of the ANOVA [$F(1, 98) = 0.816, p = 0.369$] and findings from the ANCOVA [$F(1, 97) = 1.000, p = 0.320$] were consistent in that the two groups showed no significant difference as regards increases in overall learning motivation from pre- to post-test. This is further supported by the mean plots (see Figure 14), which show that, although the intervention group increased more than the comparison group, the gap between the two lines at the post-test point was not great.

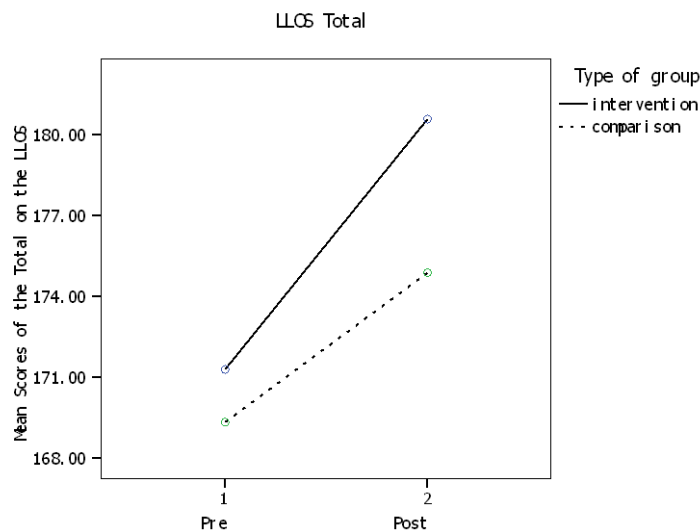


Figure 14: Pre- and post-test mean scores of the total on the LLOS

To sum up, overall findings on the LLOS indicated that the intervention group significantly improved their intrinsic motivation and identified motivation as well as their overall learning motivation from pre- to post-test, whereas the comparison group did not show significant improvements in any areas on the LLOS. However, according to the results of the ANOVA and the ANCOVA, it was only in the area of intrinsic motivation that the improvements between the two groups differed to an extent close to statistical significance. These findings were further supported by effect size statistics and mean plots. The only effect size that was above average occurred with

the intervention group in the area of intrinsic motivation. In addition, a substantial interaction effect was only found in the mean plots for intrinsic motivation.

Results from the Social Skills Scale for Chinese College English

Learners

Results from the analysis of scores on the SSS-CCEL are reported in two sections. One section is related to the eight subscales included in the SSS-CCEL. The other focuses on the items which are particularly related to the teaching and learning of English (e.g. I feel I am making progress in English learning), and excludes those which apply in general (e.g. If I work hard at something I will eventually be good at it). This separate analysis of those particular items is intended to investigate the effect of the intervention on learners' skills particularly relevant to English learning.

Results on the Eight Subscales and Total Scores

Results obtained from the analysis of scores on the eight subscales and the total scores on the SSS-CCEL are summarized in Table 4, which includes mean scores, standard deviations, Cohen's d, and p values of paired-samples t-tests within each group, the ANOVAs and the ANCOVAs, in each of the eight areas and total scores on the SSS-CCEL.

Table 4: Summary of results on the SSS-CCEL for the intervention and comparison groups

SSS-CCEL	Group	Pre/ Post	Mean	S.D.	Sig. t-test	Effect size	Sig. ANOVA	Sig. ANCOVA
Self- confidence	Int	Pre	21.577	3.226	0.001	0.470	0.160	0.210
		Post	23.250	3.875				
	Com	Pre	22.000	3.495	0.250	0.174		
		Post	22.646	3.911				
Sense of cohesion	Int	Pre	17.885	3.059	0.015	0.418	0.135	0.084
		Post	19.615	4.995				
	Com	Pre	17.458	3.701	0.180	0.139		
		Post	17.979	3.778				
Initiative in socialization	Int	Pre	21.269	3.768	0.232	0.128	0.947	0.902
		Post	21.750	3.731				
	Com	Pre	21.188	4.771	0.408	0.097		
		Post	21.625	4.266				
Being positive	Int	Pre	21.385	3.810	0.047	0.231	0.699	0.637
		Post	22.365	4.619				
	Com	Pre	21.208	5.095	0.094	0.151		
		Post	21.938	4.591				
Checking for understanding	Int	Pre	21.808	4.215	0.010	0.257	0.121	0.151
		Post	23.000	5.018				
	Com	Pre	22.333	4.582	0.812	0.027		
		Post	22.458	4.631				
Equal participation and accountability	Int	Pre	18.404	2.614	0.001	0.547	0.006	0.007
		Post	20.500	4.747				
	Com	Pre	18.771	3.985	1.000	0.000		
		Post	18.771	4.111				
Acceptance and empathy	Int	Pre	21.269	2.794	0.080	0.231	0.112	0.457
		Post	21.923	2.862				
	Com	Pre	22.708	4.332	0.551	-0.065		
		Post	22.438	3.941				
Conflict management	Int	Pre	20.731	3.861	0.148	0.130	0.236	0.376
		Post	21.231	3.812				
	Com	Pre	21.479	4.141	0.734	-0.036		
		Post	21.333	3.899				
Total	Int	Pre	164.327	19.254	0.003	0.376	0.048	0.058
		Post	173.635	29.285				
	Com	Pre	167.146	24.388	0.284	0.087		
		Post	169.188	22.333				

In the area of self-confidence, results of paired-samples t-tests showed that the intervention group [$t(51) = 3.531, p = 0.001$] increased significantly from pre- to post-test, but the comparison group [$t(47) = 1.165, p = 0.250$] did not. According to Hattie's (2009) guidelines, the effect size for the intervention group ($d = 0.470$) was above average, but not this was not the case for the comparison group ($d = 0.174$). However, both the ANOVA [$F(1, 98) = 2.003, p = 0.160$] and the ANCOVA [$F(1, 97) = 1.594, p = 0.210$] indicated that the two groups exhibited no statistically significant difference in terms of their improvements in the area of self-confidence from pre- to post-test. The mean plots for self-confidence (see Figure 15) demonstrate a clear interaction effect in favour of the intervention group.

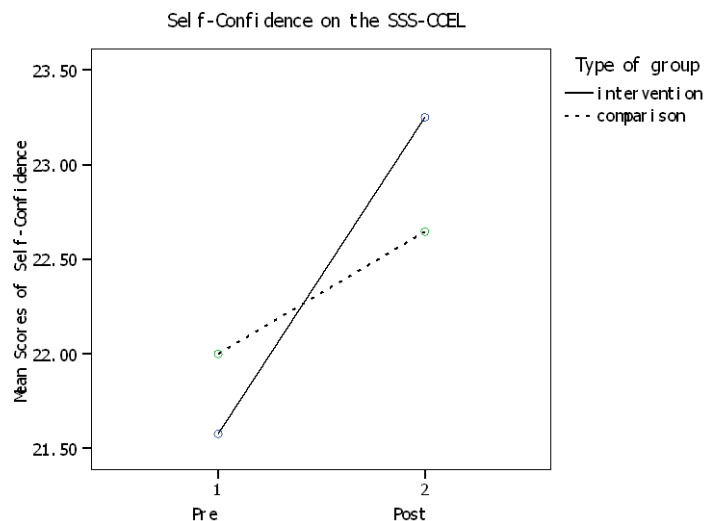


Figure 15: Pre- and post-test mean scores of self-confidence on the SSS-CCEL

Likewise, results of paired-samples t-tests regarding sense of cohesion showed that a significant increase occurred in the intervention group [$t(51) = 2.527, p = 0.015$] but not in the comparison group [$t(47) = 1.362, p = 0.180$]. The effect size was above average for the intervention group ($d = 0.418$) but not for the comparison group ($d = 0.139$). Although the ANOVA showed no significant interaction effect between time and group [$F(1, 98) = 2.275, p = 0.135$], the ANCOVA [$F(1, 97) = 3.048, p = 0.084$] indicated that, with pre-test differences held constant, post-test differences between the two groups approached statistical significance. This suggests that the intervention group achieved greater gains than the comparison group as regards sense of cohesion from pre- to post-test. The mean plots for sense of cohesion (see Figure 16) also support the assumption that there was a clear difference in the magnitude of improvements between the two groups.

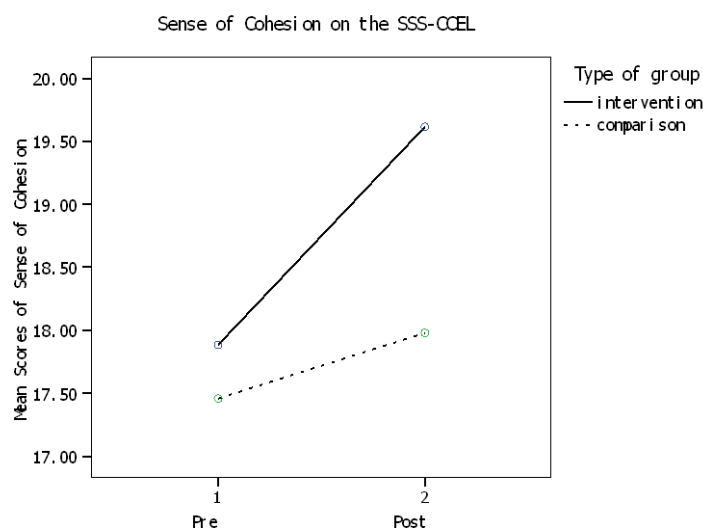


Figure 16: Pre- and post-test mean scores of sense of cohesion on the SSS-CCEL

In the area of initiative in socialization, there was no significant improvement from the pre-test to post-test for both the intervention group [$t(51) = 1.210, p = 0.232$] and the comparison group [$t(47) = 0.835, p = 0.408$]. Effect sizes for the intervention group ($d = 0.128$) and the comparison group ($d = 0.097$) were both small and well below average. Results of both the ANOVA [$F(1, 98) = 0.004, p = 0.979$] and the ANCOVA [$F(1, 97) = 0.015, p = 0.902$] showed that there was little between-group difference in enhancing initiative in socialization from pre- to post-test. This situation is once again supported by the mean plots for initiative in socialization (see Figure 17), which present two lines with very similar gradients. In addition, it should be noted that mean plots are sometimes misleading due to the use of large scales, for instance, 1: 0.2 in this case. Therefore, in spite of the visual image showing that both groups improved quite a lot in the mean plots, the actual increases were small and far from reaching statistical significance, as identified by t-tests and effect sizes.

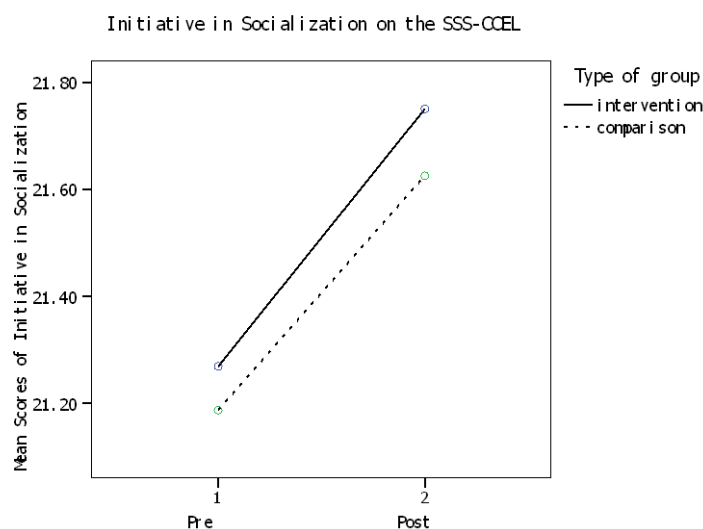


Figure 17: Pre- and post-test mean scores of initiative in socialization on the SSS-CCEL

In the area of being positive, results of t-tests indicated a significant difference between pre and post testing for the intervention group [$t(51) = 2.037, p = 0.047$] but not for the comparison group [$t(47) = 1.707, p = 0.094$]. However, effect sizes for the intervention group ($d = 0.231$) and the comparison group ($d = 0.151$) were both well below average. According to the ANOVA [$F(1, 98) = 0.151, p = 0.699$], little difference was found between the two groups as regards their improvements in this area from pre- to post-test. This finding also agreed with the result of the ANCOVA [$F(1, 97) = 0.225, p = 0.637$]. The mean plots for being positive (see Figure 18) also suggest that there was little difference in the magnitude of increases between the two groups.

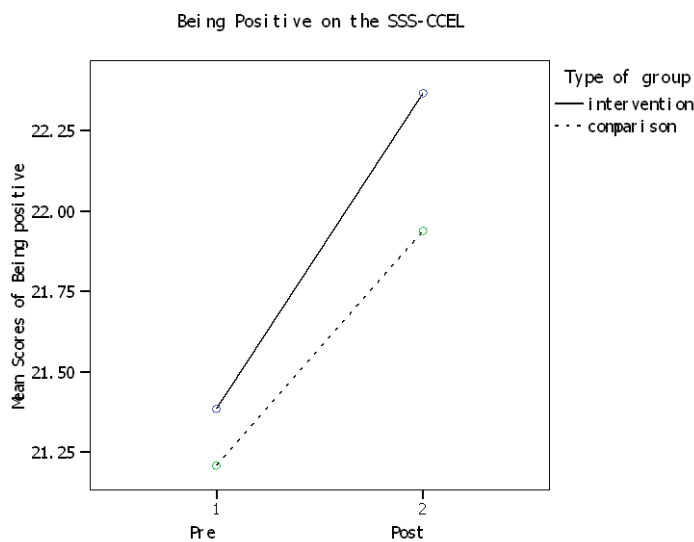


Figure 18: Pre- and post-test mean scores of being positive on the SSS-CCEL

In the area of checking for understanding, results of t-tests indicated that the intervention group [$t(51) = 2.677, p = 0.010$] improved significantly from pre- to post-test but the comparison group [$t(47) = 0.239, p = 0.812$] did not. The effect size for the intervention group ($d = 0.257$) was below average and it was very small for the comparison group ($d = 0.027$). Results of both the ANOVA [$F(1, 98) = 2.441, p = 0.121$] and the ANCOVA [$F(1, 97) = 2.099, p = 0.151$] were fairly consistent with each other. There was no statistically significant difference between the two groups in their improvements in this area. However, since both the alpha values were relatively small, it is evident that the intervention group improved more than the comparison group. The mean plots for checking for understanding (see Figure 19), display an obvious interaction effect in favour of the intervention group, with one line showing a sharp increase while the other being nearly horizontal.

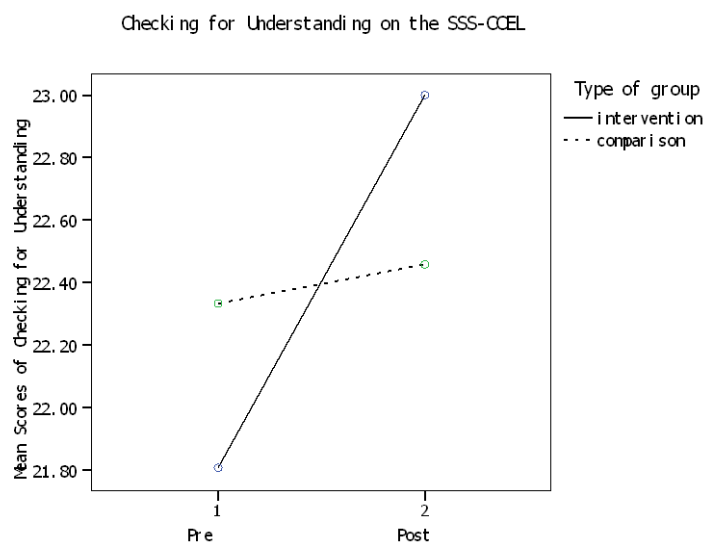


Figure 19: Pre- and post-test mean scores of checking for understanding on the SSS-CCEL

In the area of equal participation and accountability, paired-samples t-tests found statistically significant improvements for the intervention group [$t(51) = 3.382, p = 0.001$], with the effect size ($d = 0.547$) above average. The comparison group [$t(47) = 0.000, p = 1.000$] did not show any improvement. Moreover, the ANOVA showed there was a statistically significant interaction effect for time and group [$F(1, 98) = 7.888, p = 0.006$], which means that the intervention group improved significantly more than the comparison group between pre and post testing. Likewise, results of the ANCOVA [$F(1, 97) = 7.489, p = 0.007$] revealed that there was a significant post-test difference between the two groups, after adjusting for pre-test differences. This is further supported by the mean plots for equal participation and accountability (see Figure 20), which show a substantial interaction between the two lines, one staying horizontal while the other displaying a steep gradient.

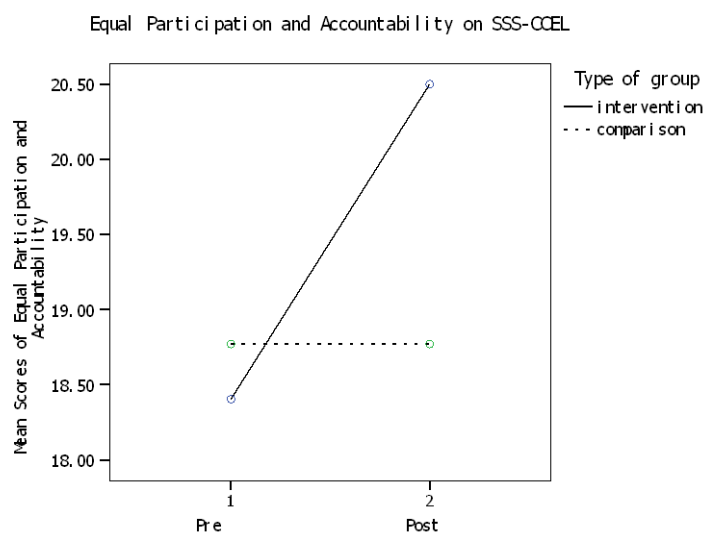


Figure 20: Pre- and post-test mean scores of equal participation and accountability on the SSS-CCEL

In the area of acceptance and empathy, the comparison group started better but showed a decline on the post-test, while in the intervention group some improvements occurred at post-test. However, results of paired-samples t-tests indicated that there was no significant difference between pre- and post-test within both the intervention group [$t(51) = 1.785, p = 0.080$] and the comparison group [$t(47) = -0.600, p = 0.551$]. Moreover, effect sizes were below average for both the intervention group ($d = 0.231$) and the comparison group ($d = -0.065$). This situation is also supported by results of both the ANOVA [$F(1, 98) = 2.565, p = 0.112$] and the ANCOVA [$F(1, 97) = 0.559, p = 0.457$], which suggest that changes which occurred in this area from pre- to post-test were not significantly different between the two groups. This is in agreement with the mean plots for acceptance and empathy (see Figure 21), which display no substantial interaction effect between the two lines, in that the comparison group remained clearly better than the intervention group on the post-test.

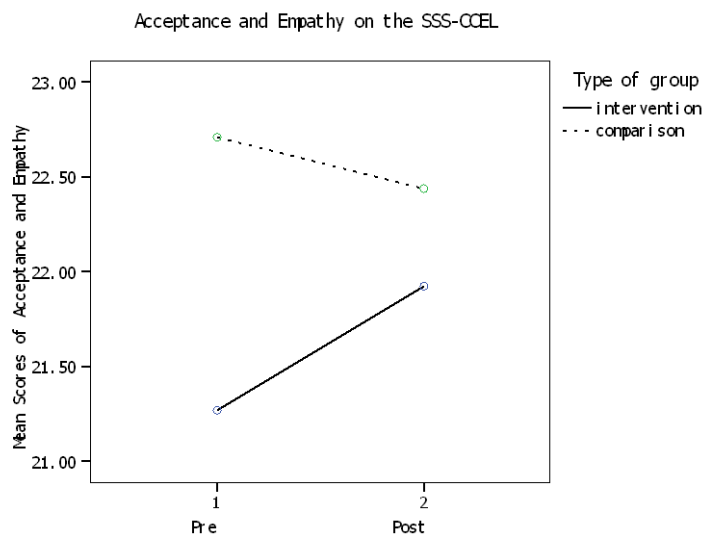


Figure 21: Pre- and post-test mean scores of acceptance and empathy on the SSS-CCEL

In the area of conflict management, the situation was quite similar to the area of acceptance and empathy; that is, the intervention group started slightly lower in this area and made some improvements on the post-test while the comparison group showed the opposite tendency. However, paired-samples t-tests indicated no significant difference from pre-test to post-test within both the intervention group [$t(51) = 1.470, p = 0.148$] and the comparison group [$t(47) = -0.341, p = 0.734$]. Effect sizes for both the intervention group ($d = 0.130$) and the comparison group ($d = -0.036$) were well below average. Also, the ANOVA [$F(1, 98) = 1.419, p = 0.236$] agreed with the ANCOVA [$F(1, 97) = 0.790, p = 0.376$] indicating that the two groups demonstrated no significant difference in their change in this area from pre- to post-test. This is further confirmed by the mean plots for conflict management (see Figure 22) which display no sign of a substantial interaction effect between the two lines, especially considering the use of a large scale of 1:0.2.

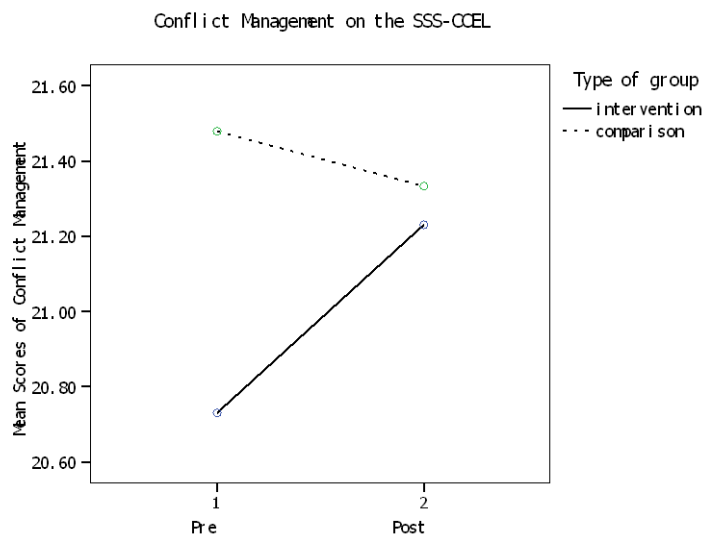


Figure 22: Pre- and post-test mean scores of conflict management on the SSS-CCEL

As for total scores on the SSS-CCEL, results of paired-samples t-tests showed that the intervention group [$t(51) = 3.085, p = 0.003$] significantly increased from pre-test to post-test while the comparison group [$t(47) = 1.084, p = 0.284$] did not. The effect size for the intervention group ($d = 0.376$) was very close to the average level, but it was very small for the comparison group ($d = 0.087$). The ANOVA [$F(1, 98) = 4.016, p = 0.048$] showed a statistically significant interaction effect for time and group although the alpha value was marginal. Consistent with these findings, the ANCOVA [$F(1, 97) = 3.666, p = 0.058$] also generated a marginal alpha value. Thus, these results are indicative of a trend towards greater improvements in the intervention group. This is further supported by inspection of the mean plots for overall social skills (see Figure 23), which show the intervention group improved considerably more than the comparison group because there is a substantial interaction between the two lines.

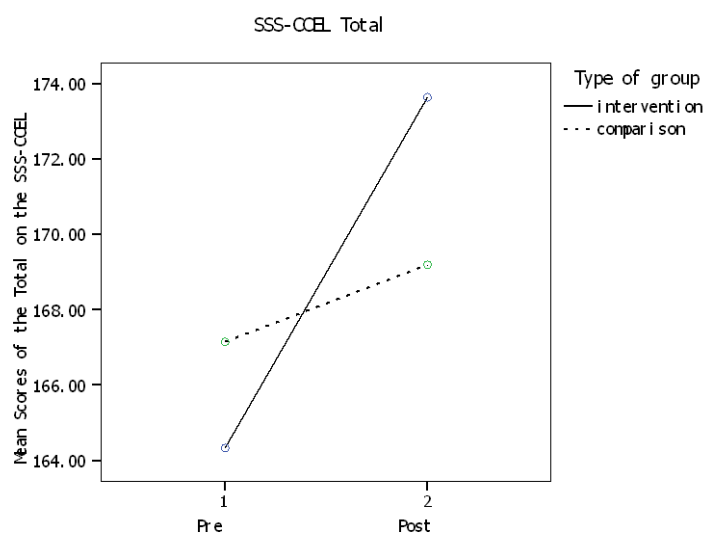


Figure 23: Pre- and post-test mean scores of the total on the SSS-CCEL

To sum up, overall findings on the SSS-CCEL indicated that the intervention group improved significantly in their overall social skills while the comparison group did not. Among the eight areas, there were five—self-confidence, sense of cohesion, being positive, checking for understanding, and equal participation and accountability—where the intervention group made significant improvements between pre and post testing, while the comparison group did not

improve significantly in any of the areas. The effect size statistics showed that there were only three areas—self-confidence, sense of cohesion, as well as equal participation and accountability—where the magnitude of improvement within the intervention group was above the average level of 0.4. However, results of the ANOVA and the ANCOVA indicated that it was only in the area of equal participation and accountability that the intervention group improved significantly more than the comparison group between pre and post testing. It is also notable that between-group differences in terms of their improvements in total scores on the SSS-CCEL were statistically significant at a marginal level.

Results from Items Focusing on Learning of English

Among the 40 items of the SSS-CCEL, there are a total of nine items which contain the particular word “English” and are specifically related to the learning of English, for instance, “I am doing a good job of English language learning”. In this thesis, these nine items are called focus items, and the specific skills relevant to these focus items are thus termed focus skills.

Results obtained from the analysis of data based on these nine items are summarized in Table 5, which includes mean scores, standard deviations, Cohen’s d, and p values of paired-samples t-tests within each group, the ANOVAs and the ANCOVAs, relating to each item and total scores on the nine items.

Table 5: Summary of results from items focusing on learning of English for the intervention and comparison groups

Focus items	Group	Pre/ Post	Mean	S.D.	Sig. t-test	Effect size	Sig. ANOVA	Sig. ANCOVA
Item 3	Int	Pre	2.962	0.885	0.001	0.492	0.222	0.407
		Post	3.520	1.336				
	Com	Pre	3.229	1.016	0.051	0.282		
		Post	3.521	1.052				
Item 4	Int	Pre	3.269	1.087	0.000	0.881	0.067	0.153
		Post	4.212	1.054				
	Com	Pre	3.521	1.052	0.001	0.504		
		Post	4.063	1.099				
Item 7	Int	Pre	2.981	0.980	0.058	0.339	0.149	0.126
		Post	3.365	1.268				
	Com	Pre	3.000	0.923	0.728	0.042		
		Post	3.042	1.051				
Item 10	Int	Pre	3.885	0.832	0.049	0.313	0.022	0.006
		Post	4.231	1.323				
	Com	Pre	3.729	1.317	0.221	-0.137		
		Post	3.563	1.090				
Item 15	Int	Pre	5.231	1.113	0.015	0.343	0.000	0.001
		Post	5.596	1.015				
	Com	Pre	5.667	0.930	0.003	-0.527		
		Post	5.104	1.189				
Item 25	Int	Pre	4.173	1.200	0.002	0.386	0.146	0.261
		Post	4.673	1.382				
	Com	Pre	4.438	1.183	0.448	0.115		
		Post	4.583	1.334				
Item 26	Int	Pre	3.789	0.667	0.001	0.785	0.007	0.000
		Post	4.500	1.094				
	Com	Pre	3.625	0.815	0.767	0.047		
		Post	3.667	0.975				
Item 27	Int	Pre	3.712	1.016	0.001	0.403	0.061	0.057
		Post	4.192	1.344				
	Com	Pre	3.771	1.448	0.824	0.030		
		Post	3.813	1.394				
Item 35	Int	Pre	3.365	0.841	0.000	0.838	0.000	0.001
		Post	4.135	0.991				
	Com	Pre	4.083	1.252	0.211	-0.140		
		Post	3.896	1.148				
Total	Int	Pre	33.365	3.742	0.000	0.792	0.000	0.000
		Post	38.423	8.223				
	Com	Pre	35.063	5.220	0.714	0.034		
		Post	35.250	5.617				

Item 3 is “I am doing a good job of English language learning”. Results of paired-samples t-tests showed that the intervention group [$t(51) = 3.513, p = 0.001$] increased significantly from pre-test to post-test, and the comparison group [$t(47) = 2.001, p = 0.051$] also increased an amount which was marginally significant. According to Hattie’s (2009) guidelines, the effect size for the intervention group ($d = 0.492$) was above average, but it was not for the comparison group ($d = 0.282$). Although the intervention group apparently enhanced their confidence in English learning more than the comparison group, both the ANOVA [$F(1, 98) = 1.509, p = 0.222$] and the ANCOVA [$F(1, 97) = 0.695, p = 0.407$] showed no statistical significant difference between the groups in improvements. This situation is also reflected by the mean plots (see figure 24), from which it can be seen that the intervention group started at a lower level but achieved a similar level to the comparison group on the post-test; however the interaction effect was not substantial.

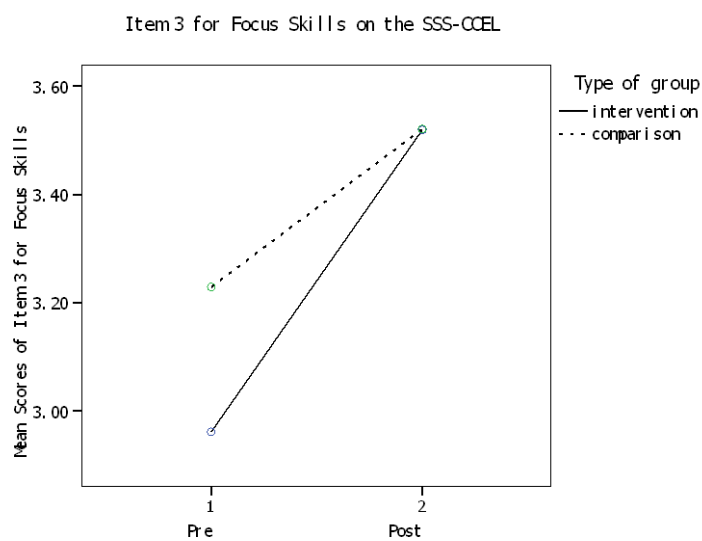


Figure 24: Pre- and post-test mean scores of item 3 for focus skills on the SSS-CCEL

Item 4 is “I feel I am making progress in English language learning”. Results of t-tests indicated statistically significant improvements between the two time periods for both the intervention group [$t(51) = 6.327, p < 0.0005$] and the comparison group [$t(47) = 3.441, p = 0.001$]. Effect sizes for the intervention group ($d = 0.881$) and comparison group ($d = 0.504$) were both above average, but there was a noticeable difference of 0.377 in between. Although the interaction effect [$F(1, 98) = 3.422, p = 0.067$] was not statistically significant, it suggests a trend of greater gains in the intervention group. The ANCOVA [$F(1, 97) = 2.074, p = 0.153$] showed no statistical significant between-group difference in their improvements at post-test, after controlling for pre-test differences. The mean plots (see figure 25) show that although both groups improved, the intervention group clearly improved more especially considering the use of a relatively big scale unit of 0.25 in this case.

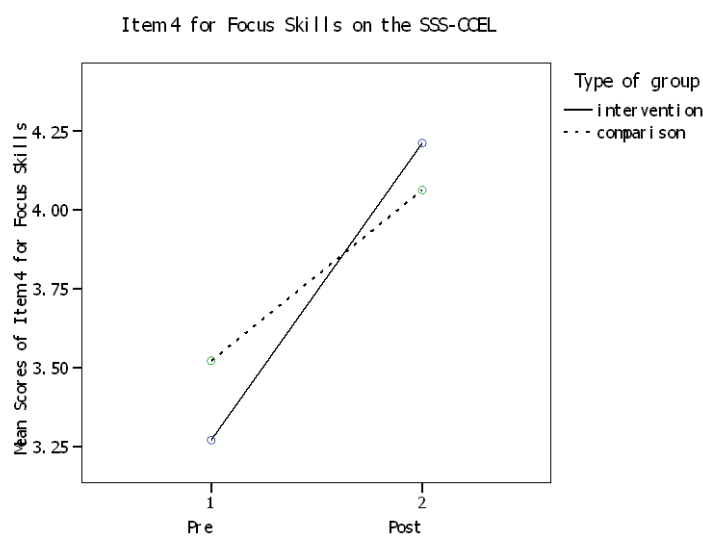


Figure 25: Pre- and post-test mean scores of item 4 for focus skills on the SSS-CCEL

Item 7 is “I consider peer support indispensable to my English language learning and success”. According to t-tests, although both the intervention group [$t(51) = 1.939, p = 0.058$] and the comparison group [$t(47) = 0.728, p = 0.350$] did not significantly improve, the alpha value for the intervention group was very marginal. This was in accordance with Cohen’s d statistics, which showed that the effect size for the intervention group ($d = 0.339$) approached the average level, but was extremely small for the comparison ($d = 0.042$). However, the findings from both the ANOVA [$F(1, 98) = 2.112, p = 0.149$] and the ANCOVA [$F(1, 97) = 2.380, p = 0.126$] indicated no significant between-group difference as regards improvements in this skill from pre- to post-test. According to the mean plots (see figure 26), there was a clear interaction effect between the two lines, but a check of the scale unit, which is only 0.1 point, reveals that the actual magnitude of interaction is not as sizable as is shown visually.

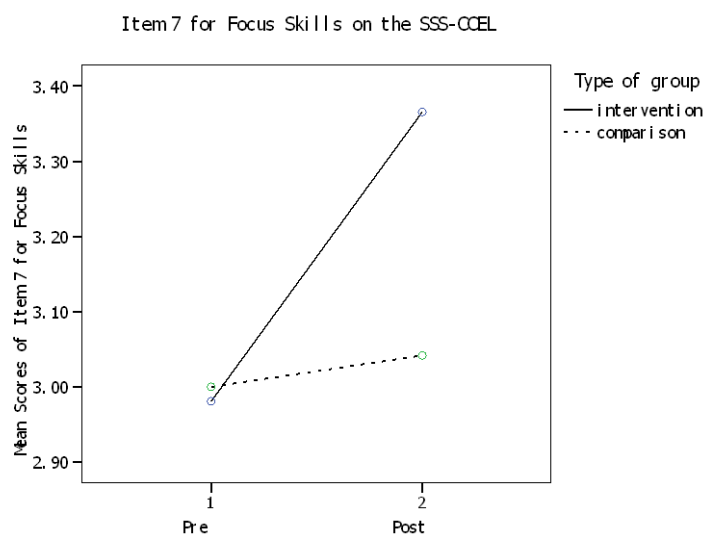


Figure 26: Pre- and post-test mean scores of item 7 for focus skills on the SSS-CCEL

Item 10 is “I offer teammates support and assistance as much as I can, so that they do their best in English”. The intervention group [$t(51) = 2.021, p = 0.049$] exhibited a marginally significant increase on this item between the two time periods, while the comparison group [$t(47) = -1.241, p = 0.221$] demonstrated a decrease. The effect size for the intervention group ($d = 0.313$) approached the average level, while it was negative for the comparison group ($d = -0.137$). The result of the ANOVA [$F(1, 98) = 5.431, p = 0.022$] was consistent with the finding from the ANCOVA [$F(1, 97) = 7.824, p = 0.006$], which indicated that the intervention group significantly outperformed the comparison group regarding the improvements on this item. This is supported by the mean plots (see figure 27) in which the two lines present a substantial interaction by showing very different gradients.

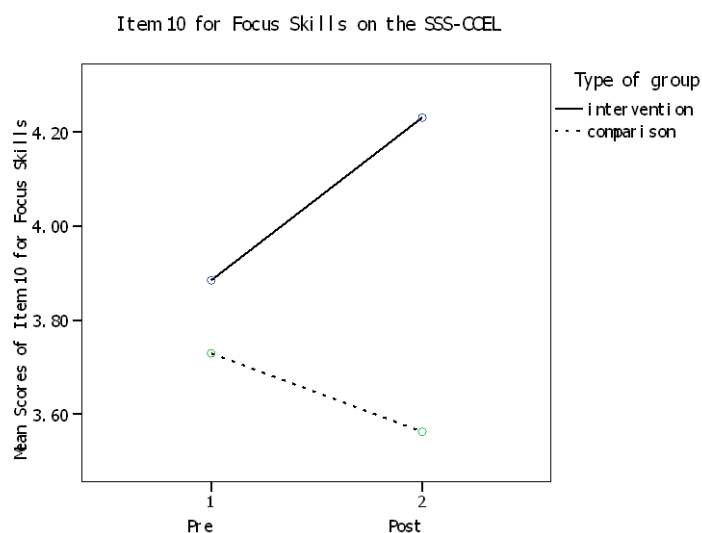


Figure 27: Pre- and post-test mean scores of item 10 for focus skills on the SSS-CCEL

Item 15 is “I feel stressed and uncomfortable when working with others in English classes”. Results of t-tests indicated that the intervention group [$t(51) = 2.513, p = 0.015$] significantly improved from pre to post-test, while the comparison group [$t(47) = -3.110, p = 0.003$] showed the opposite tendency. The effect size for the intervention group ($d = 0.343$) was close to the average level, whereas it was negative for the comparison group ($d = -0.527$) which indicated a striking decrease. Understandably, the ANOVA [$F(1, 98) = 16.216, p < 0.0005$] and the ANCOVA [$F(1, 97) = 11.135, p = 0.001$] both indicated that, compared with the comparison group, the intervention group significantly improved between pre and post testing. This finding was once again supported by the mean plots (see figure 28), which depict two lines dramatically interacting with each other.

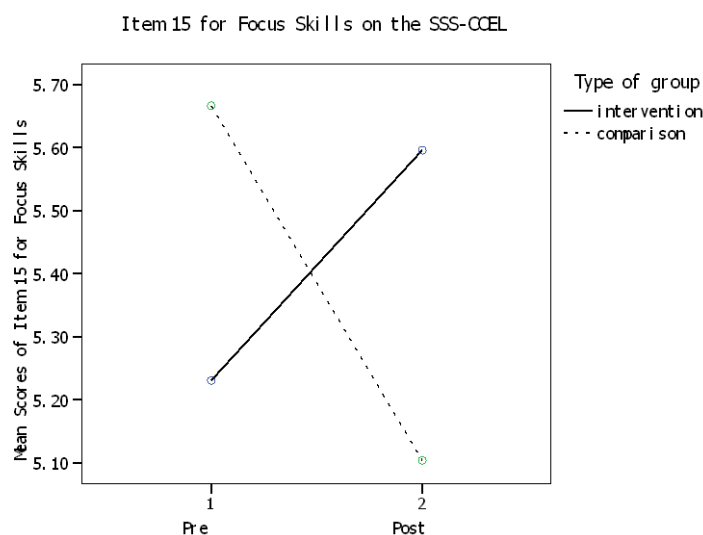


Figure 28: Pre- and post-test mean scores of item 15 for focus skills on the SSS-CCEL

Item 25 is “I often use examples to make myself understood in English classes”. According to the results of t-tests, the intervention group [$t(51) = 3.297, p = 0.002$] significantly improved from pre-test to post-test, while the comparison group [$t(47) = 0.765, p = 0.448$] did not. The effect size for the intervention group ($d = 0.386$) was close to the average level of 0.4, while it was small for the comparison group ($d = 0.115$). The results of both ANOVA [$F(1, 98) = 2.146, p = 0.146$] and ANCOVA [$F(1, 97) = 1.279, p = 0.261$] indicated that the two groups demonstrated no statistically significant difference in terms of their improvements on this item. This situation is also reflected by the mean plots (see figure 29), that is, although the intervention group improved more than the comparison group, the interaction between the two lines is not substantial, especially considering the use of a small scale unit of 0.1 in this case.

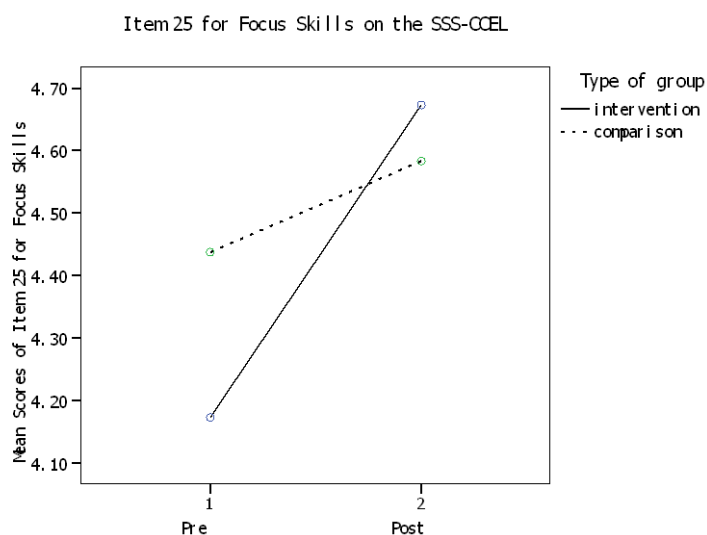


Figure 29: Pre- and post-test mean scores of item 25 for focus skills on the SSS-CCEL

Item 26 is “I have a clear picture of my personal role in English teamwork and participate actively”. The results of t-tests indicated the intervention group [$t(51) = 3.689, p = 0.001$] significantly improved from pre-test to post-test, while the comparison group [$t(47) = 0.299, p = 0.767$] did not change much. The effect size for the intervention group ($d = 0.785$) was well above the average level, whereas it was very small for the comparison group ($d = 0.047$). The ANOVA [$F(1, 98) = 7.698, p = 0.007$] showed a significant interaction effect between time and group. This finding was in accordance with that from the ANCOVA [$F(1, 97) = 14.950, p < 0.0005$], which showed a significant post-test difference between groups after adjusting for pre-test differences. This is further confirmed by the mean plots (see figure 30), in which the line for the comparison group stays almost horizontal from pre- to post-test, while the line for the intervention group shows a striking increase.

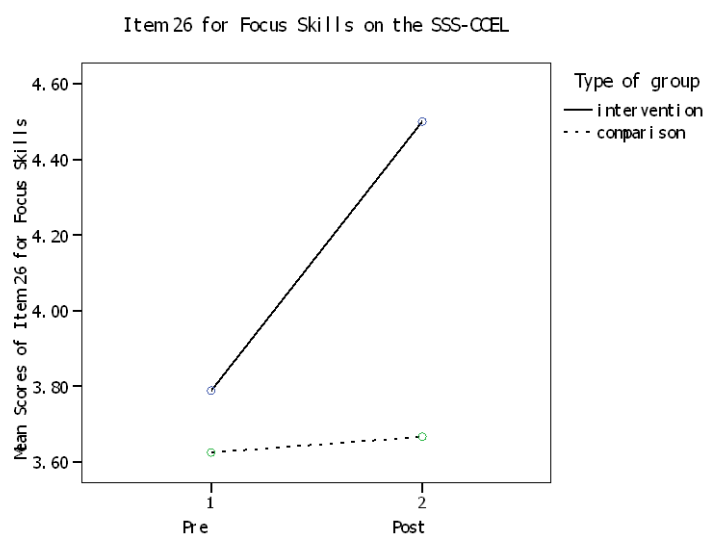


Figure 30: Pre- and post-test mean scores of item 26 for focus skills on the SSS-CCEL

Item 27 is “I tend to keep silent in English classes, except when I am called to answer questions”. According to t-tests, the intervention group [$t(51) = 3.401, p = 0.001$] gained significant improvements at post-test, while the comparison group [$t(47) = 0.224, p = 0.824$] changed little. The effect size statistics showed that the magnitude of improvements in the intervention group ($d = 0.403$) was slightly above average, while it was almost negligible in the comparison group ($d = 0.030$). Although the results from the ANOVA [$F(1, 98) = 3.603, p = 0.061$] and the ANCOVA [$F(1, 97) = 3.714, p = 0.057$] did not show a statistically significant difference between the two groups, the alpha values were both very marginal and close to the significant level. This suggested that the intervention group obviously improved more than the comparison group on this item, but the between-group difference did not achieve the statistical significance yet. This finding is clearly reflected by the mean plots (see figure 31), which show a clear interaction between the two lines.

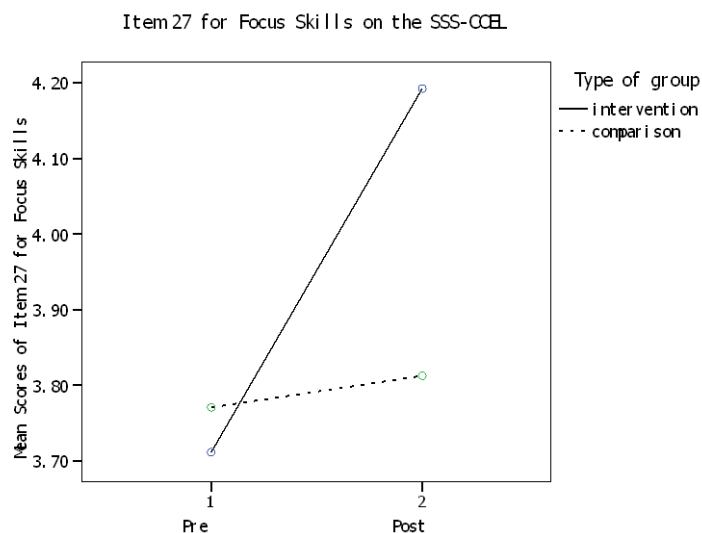


Figure 31: Pre- and post-test mean scores of item 27 for focus skills on the SSS-CCEL

Item 35 is “I believe everybody of different English levels can make a contribution to the completion of group tasks”. The t-tests showed that the intervention group [$t(51) = 5.430, p < 0.0005$] significantly improved from pre-test to post-test, while the comparison group [$t(47) = -1.268, p = 0.211$] decreased slightly. The effect size for the intervention group ($d = 0.838$) was large and much higher than the average cut-off point of 0.4, while it was negative and small for the comparison group ($d = -0.140$). The ANOVA [$F(1, 98) = 21.834, p < 0.0005$] and the ANCOVA [$F(1, 97) = 12.717, p = 0.001$] were consistent with each other showing that, there was a statistically significant difference between the two groups as regards improvements from pre to post-post. The mean plots (see figure 32) provide further supporting evidence for this situation through a substantial interaction between the two lines, one greatly rising while the other slightly declining.

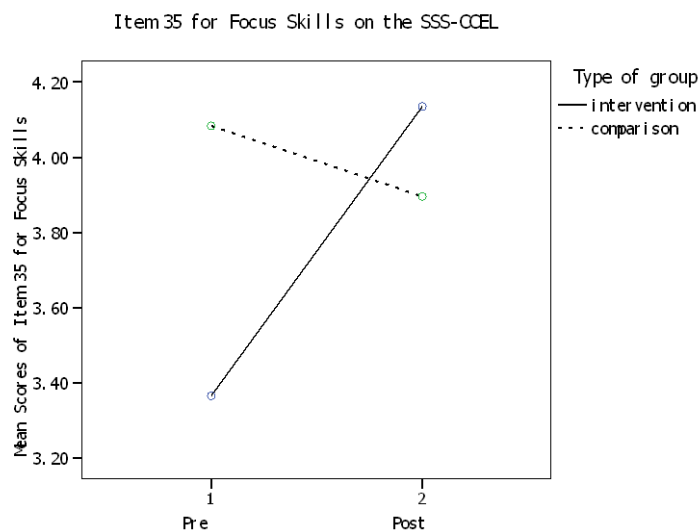


Figure 32: Pre- and post-test mean scores of item 35 for focus skills on the SSS-CCEL

As for total scores of the nine focus items, results of paired-samples t-tests showed that students in the intervention group [$t(51) = 5.090, p < 0.0005$] significantly improved their ratings of these skills from pre- to post-test, while those in the comparison group [$t(47) = 0.368, p = 0.714$] changed little. The effect size for the intervention group ($d = 0.792$) was large and well above average, while it was very small for the comparison group ($d = 0.034$). Results of the ANOVA [$F(1, 98) = 18.107, p < 0.0005$] and the ANCOVA [$F(1, 97) = 16.547, p < 0.0005$] agreed with the finding that the intervention group improved significantly more than the comparison group between pre and post testing. This is further confirmed by the mean plots for overall focus skills (see figure 33), which shows a substantial interaction between the two lines, one staying nearly horizontal from pre- to post-test while the other rising sharply.

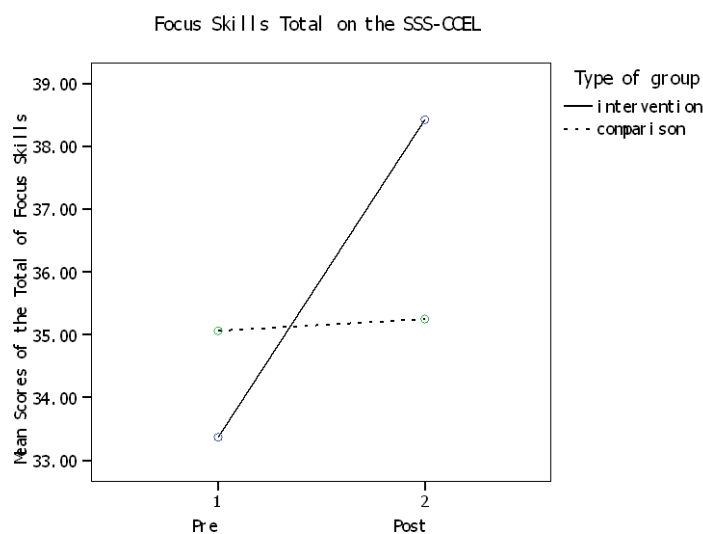


Figure 33: Pre- and post-test mean scores of the total for focus skills on the SSS-CCEL

To sum up, the overall findings on these nine items indicated a substantial increase in favour of the intervention group, who reported significantly better performance at post-test than at pre-test regarding each of related aspects, as well as the overall focus skills. In contrast, the comparison group showed a statistically significant increase in only one aspect (relating to item 4 on progress-making in English) and a significant decrease in one aspect (relating to item 15 on their

feelings of comfort when working in teams), and thus showed little improvement in their ratings of overall focus skills. Moreover, effect sizes for the intervention group were above average in five aspects (relating to item 3, 4, 26, 27, and 35) as well as in the overall focus skills, while for the comparison group there was only one aspect (relating to item 4) which showed this desirable magnitude of improvement. Results of the ANOVA and the ANCOVA indicated that the intervention group improved significantly more than the comparison group in four specific aspects (relating to item 10, 15, 26, and 35) as well as in the overall focus skills. It is also noted that the alpha values of the ANOVA and the ANCOVA approached statistical significance in a couple of aspects (relating to item 4 and 27).

CHAPTER FIVE: DISCUSSION

Overview

This chapter discusses findings with regard to Chinese tertiary learners' English proficiency, motivation for learning English, and their social skills. Statistical analyses indicate that there were a number of areas on which the intervention group, taught with the cooperative learning (CL) approach, substantially outperformed the comparison group instructed by traditional methods. These areas include listening, speaking and reading, as well as their overall English proficiency. With regard to motivation, the CL approach had greater impact than traditional instruction on intrinsic motivation. Regarding social skills, substantial between-group differences were found in the area of equal participation and accountability, as well as in overall social skills, in favour of the intervention group. In addition, the intervention group was found to have made substantially more improvements than the comparison group in specific social skills directly related to the learning of English language. This chapter includes a discussion about findings of the current study in relation to the findings of previous research. It goes on to discuss implications for the practice of EFL teaching, with a focus on the challenges of using CL in Chinese tertiary institutions. At the end of this chapter, the major contributions and limitations of the study are discussed, and recommendations for future research are suggested.

English Language Proficiency

One of the major findings of this research is that the intervention group, taught using the CL approach, made substantially greater gains than the comparison group in their overall English proficiency. As for the five specific areas relating to English proficiency, CL was found to be

considerably more effective than traditional methods in teaching speaking and listening. In the area of reading, it was also found that the intervention group clearly improved more than the comparison group, but between-group differences were marginal. However, the two teaching approaches were found not to differ in their effectiveness in the areas of writing and vocabulary. These findings and the possible reasons for them are discussed below.

The finding that the CL approach is more effective than traditional instruction in improving learners' English proficiency supports previous findings regarding its role in enhancing academic achievement (Brown & Thomson, 2000; Johnson et al., 1998; Kagan, 1994; Slavin, 1995). Actually, a good number of large-scale meta-analyses (e.g. Hattie, 2009; Johnson et al., 1981; Slavin, 1995) have already highlighted the strengths of CL over the traditional whole-class instruction in improving learners' academic proficiency, which is considered one of the major positive outcomes of CL. Findings of the study reported in this thesis also support the view that CL is more effective in teaching language because of its efficacy in satisfying the communicative nature of language acquisition through the maximum use of promotive peer interaction in a positively-interdependent and non-threatening environment (Dörnyei, 1997; High, 1993; Holt, 1993; Jacobs & Goh, 2007; Kagan & McGroarty, 1993). In addition, the findings of this study support the perspective that well-structured CL teamwork can be adapted to fit in the foreign language teaching context (Flowerdew, 1998; Jacobs & McCafferty, 2006; Magee & Jacobs, 2001).

The largest difference between the impact of the CL approach and that of traditional instruction was in the area of increasing students' speaking competence. This finding supports the widely-accepted view that CL facilitates the development of verbal skills (Dörnyei, 1997; Jacobs and Goh, 2007; Jacobs and McCafferty, 2006; Kagan, 1994; McGroarty, 1993). It also supports findings from a previous study with Chinese learners which reported the CL approach to be significantly more effective than traditional instruction in improving learners' speaking competence (Chen, 2005). It is considered that three components of CL contributed to its effectiveness in improving students' speaking ability in the current research. First, the quantity of communication in English in the class was considerably increased in the intervention group. This

is because, in contrast with the comparison group which was mainly exposed to whole-class instruction, the intervention group frequently used CL activities which involved a considerable amount of peer interaction. With the whole-class format used in traditional instruction only one student was allowed to speak at a time, but with the foursome group work used in the intervention group, 13 of the 52 students could talk simultaneously. Also the supportive and non-threatening learning environment created through the use of CL appeared to reduce the intervention group's anxiety and made them feel safe and enthusiastic enough to take risks in order to speak English. Second, the quality of language production was improved in the intervention group by frequently working on meaningful real-life team tasks, which also stimulated students to perform more language functions rather than merely rote-learning language forms. For instance, while working on team tasks, they needed to use a variety of language functions, such as asking for repetition or clarification, checking for comprehension of listeners, paraphrasing for easier understanding, and complimenting or encouraging each other in order to build positive team relations. This contrasted with the language output produced by the comparison group, which was typically inauthentic and inadequate. This was probably because students in the comparison group had fewer opportunities for real-life communication with peers when exposed to whole-class instruction, so that mostly their language practice was related to particular grammatical or lexical items selected from the text. Third, the intervention group gained a great deal of positive experience with English through a wealth of promotive peer interaction in the supportive learning environment. This experience of using English as a communication tool, instead of studying it as purely linguistic knowledge, enhanced the intervention group's interest in speaking and trying out the language.

The finding that CL was substantially better than traditional teaching in enhancing students' listening competence is consistent with the findings of two previous studies (Bejarano, 1987; Chen 2005), which reported that students in CL classes made significantly greater gains than their peers in traditional classes on listening comprehension. Improvements in listening and speaking should go hand in hand because peer interaction is two-way communication in which students are shifted from the role of listeners to speakers in turn (Long & Porter, 1985; McGroarty, 1993). In the current research, the finding that the intervention group did not make as much progress in listening

as in speaking could be explained by three factors. First, peer interaction included a lot of meaning negotiation and modified speech since students tailored language forms and content to a level that their peer listeners could understand. So a substantial difference existed between peer talk and the content of CET listening comprehension test materials in terms of the degree of difficulty. Second, students were allowed to use a small amount of Chinese when they worked on their team tasks, before presenting in English. It was noticed that students tended to use Chinese when finding it difficult to get complex ideas across. This reduced the challenge of practice for listeners to some extent. Thirdly, listening on the CET involved two parts: one was the multiple-choice test for listening comprehension, and the other was spot dictation, which required students to fill in blanks according to what they had heard from a recording. Spot dictation, as a mixed test of listening, vocabulary and grammar, also emphasizes accuracy of language forms and spelling. So students needed not only to be able to understand the listening materials but also to be able to take down the missing words in terms of correct spelling and grammatical forms, whereas accuracy of spelling and grammar was not particularly emphasized in the CL classroom. Thus, the intervention group might not outperform the comparison group in completing spot dictation testing items.

This research also found that there was a substantial difference, in favour of the CL approach, regarding students' reading competence, although this difference was statistically marginal. This finding is interesting in relation to the findings of four previous studies (Bejarano, 1987; Chen, 2005; Ghaith, 2003; Waugh et al., 2005) which have produced somewhat conflicting results. Two studies in Lebanon (Ghaith, 2003) and Thailand (Waugh et al., 2005) reported that CL greatly facilitated improvements in learners' reading comprehension as compared with traditional direct instruction, while the other two studies in Israel (Bejarano, 1987) and Taiwan (Chen, 2005) stated that no substantial difference was found between the two methods in teaching reading. However, the discrepancy in findings may be due to some variation in the level of exposure to the intervention in different studies. Bejarano (1987, p. 497) noted in her study that "the students in the small-group classes were neither given tasks designed to develop reading strategies, nor were they encouraged to do so in any way". Likewise, Chen (2005) found in his study that students often read silently on their own without interaction when processing reading materials, which

undoubtedly resulted in weakening the intervention. In contrast, in the current research, cooperative team discussion was frequently used within the intervention group when reading passages were taught. The questions or topics for discussion were usually open-ended and associated with summarizing, inferring, commenting, analyzing, reasoning, negotiating and decision-making, which not only offered students opportunities to develop a deeper understanding of texts through peer interaction and consultation but also generated a substantial exchange of reading strategies (e.g. strategies related to scanning for bits of particular information, skimming for a general idea, finding topic sentences, making predictions and inferences) which are often involved in processing the reading comprehension materials on the CET. All this helped transfer a traditionally passive activity which mainly involved teacher talk and explanations into active learning, which was therefore likely to be more productive and beneficial for students. Therefore, in comparison with the direct instruction mainly used in traditional classroom, the CL reading tasks used with the intervention group in the current research facilitated the development of learners' reading skills, and further benefited students' speed and accuracy in processing reading materials on the reading comprehension test.

The finding that there was little difference between the intervention and comparison groups in vocabulary improvement is consistent with the finding of a previous study (Bejarano, 1987), which found little difference between the CL method of Group Investigation (GI) and traditional whole-class instruction in teaching vocabulary in Israeli EFL classrooms. However, in the same study, Bejarano (1987) also found that Student-Team Achievement-Division (STAD) was substantially superior to GI and whole-class instruction in improving students' vocabulary scores. The finding of the current research also contrasts with Gömleksiz's (2007) finding that CL significantly enhanced students' vocabulary knowledge and grammar use as compared with conventional instruction. These contrasting results were possibly due to the fact that the current research laid emphasis on the fluency of language use while the accuracy of vocabulary and grammar usage was stressed in the other studies. This explanation is in accordance with Bejarano's (1987, p. 497) interpretation that the cause for different findings regarding STAD and GI in her study lay in the basic difference in instructional objectives between the two methods,

that is, STAD was intended for “the learning of structural language material presented initially by the teacher” with emphasis on the accuracy of language usage while GI tasks were designed for “global use of language in real communication” with emphasis on fluency. So the intended learning objective of CL team tasks formed a very important factor in influencing the learning outcomes. In addition, frequent peer interaction in English might enable the intervention group to build up a wider active vocabulary which they could put into actual use. However, since students in the intervention group more often used easy and informal vocabulary to make themselves understood when communicating with peers, it is understandable that they might not outperform the comparison group in accurate use of difficult words, since the vocabulary test of the CET focused more on difficult academic vocabulary.

Likewise, little difference was found between the two teaching approaches in improving students’ writing scores. This may be related to the fact that the intervention group was no better than the comparison group in terms of accuracy of grammar usage, spelling and other aspects of writing, although they were found to have written more within the given time and also outperformed the comparison group in the ability to get ideas across. However, they tended to make similar writing mistakes to the comparison group, including misspellings, use of incomplete sentences, and wrong use of punctuation and verb forms, which were critical factors in assessing writing on the CET, although mostly inconsequential in oral communication. The finding from the writing test in the current research shows some inconsistency with McGuire’s (1992) finding with Japanese students, which was in favour of the CL approach in terms of both the length of writing and error occurrences in writing. But one critical point that cannot be neglected in McGuire’s study is that students in CL and traditional settings completed writing tasks in different ways—the CL students worked together in supportive groups while their peers in the traditional classroom wrote individually all by themselves. This involvement of peer support and peer correction in the CL classroom reported by McGuire may be a major reason why the students in CL groups registered significantly greater gains than their peers in the traditional setting, especially in terms of occurrences of writing errors.

Motivation for Learning English

The major finding on the Language Learning Orientations Scale (LLOS) is that the intervention group, which was taught using a CL approach, improved intrinsic motivation substantially more than the comparison group instructed by traditional methods, although the difference was only statistically significant at a marginal level. However, there were no statistically significant differences between the two groups in the other five aspects of motivation, as well as in overall motivation towards English language learning. These findings and the possible reasons for them are discussed below.

The finding that the CL approach was superior to traditional instruction in enhancing learners' intrinsic motivation supports the widely accepted view that CL generates high intrinsic motivation (Brown & Thomson, 2000; Dörnyei, 1997; Holt, 1993; Jacobs & Goh, 2007; Jacobs, Power & Inn, 2002; Johnson & Johnson, 2003; Johnson et al., 1998; Kagan, 1994; Kagan & Kagan, 2009; McCafferty et al., 2006; Slavin, 1995). This also supports results from a previous study which reported the superiority of CL over traditional teaching in improving EFL students' intrinsic motivation in the Taiwanese educational setting (Chen, 2005). Supportive evidence was also revealed in other studies (Gömleksiz, 2007; Sachs et al., 2003; Waugh et al., 2005) which indicated that, in comparison with traditional teaching, CL was more likely to improve students' positive attitudes towards English learning, which may facilitate students' interest and intrinsic motivation to learn. The positive link between CL and intrinsic motivation may be primarily attributed to the ability of CL to facilitate a supportive and non-threatening learning atmosphere where students find it fun and enjoyable to learn and thus are intrinsically motivated to achieve goals.

In this research, there are five additional possible reasons for students' increased intrinsic motivation. First, students had opportunities to use English as a communicative tool when working on team tasks. When students found they could successfully use the English language to complete a task or do something real and meaningful, they experienced positive feelings and a high level of

satisfaction from the outcomes of their learning. This experience was accompanied by a sense of accomplishment which further stimulated their desire and interests to learn and try out the language. Second, students were allowed more ownership and control over their learning in this research, for instance, by choosing their topics, sub-tasks or ways of completing teamwork, assigning individual accountability to team members, as well as assessing their learning outcomes through self and peer grading. Empowering students and increasing their autonomy in learning were probably also conducive to the enhancement of intrinsic motivation (Baloche, 1998). Third, intrinsic motivation tends to result from formative assessment and meaningful feedback relevant to the extent to which students competently complete their current tasks (Boud et al., 2001; Johnson & Johnson, 2003). In the current research, meaningful feedback was provided by peers and the teacher immediately after selected teams had presented their work. Through feedback from others, students got a clear picture about what their strengths and weaknesses were and how they could improve their work, which meant that they felt their work and effort were valued and recognized by the teacher and their peers. This facilitated their intrinsic motivation for setting up new attainable goals and getting engaged in learning tasks. Fourth, it is believed that, when students realize that their personal achievement may benefit others, they are intrinsically motivated to accomplish more (Gagne & Deci, 2005; Noels et al., 2000). In this research, students in the CL classroom were clear that everyone stood a chance of representing their home team and their success would benefit other teammates. This facilitated teammates' enthusiasm for positive interpersonal support and further promoted students' intrinsic motivation to improve and achieve. Fifth, the technique of improvement points used in the interventions encouraged students to attempt to surpass themselves rather than compete with others (Slavin, 1995). Through this technique, students of different academic levels could get access to equal opportunities for success. This would have generated more enjoyment and greater sense of achievement for students, and hence they became more intrinsically motivated to commit to learning.

However, the intervention group taught with the CL approach did not show substantial increases in any of the four subtypes of extrinsic motivation—integrated motivation, identified motivation, introjected motivation, and external motivation—compared with the comparison

group taught in the traditional classroom. The major cause of the non-significant between-group differences in integrated motivation may lie in the fact that the intervention group, compared with the comparison group, was not provided with extra impetus to become integrated in the target culture or community. This is because students in both the CL and the traditional classroom were using exactly the same textbooks and instructional materials, as well as teaching resources and objectives. So it is understandable that the intervention group did not have advantages over the comparison group in getting access to the target culture or having personal communication with native speakers of English. It is assumed that interventions involving more direct contact with native speakers enable students to have better understanding of the target culture, which is likely to spur their desire to be part of it and thus increase their integrated motivation (Gagne & Deci, 2005; Noels et al., 2000). This particular spur for integrated motivation is always rare in the context of foreign language learning, either in the CL or traditional classroom.

The use of the same learning materials and teaching resources with both groups could also partly explain why the two groups did not differ much from each other as regards their increases in identified motivation. It is assumed that CL students' successful experiences in using the target language to complete team tasks may provide them with positive feelings about the learning of the language that facilitates students' perceptions of being the kind of person who can speak the language (Holt, 1993; Jacobs & Goh, 2007; McCafferty et al., 2006). These perceptions of students may strengthen the association between the learning of language and their identities. This assumption is supported by the finding in this research that the intervention group improved their identified motivation more than the comparison group from pre- to post-test, although the between-group difference did not reach a statistical significance. This suggests that the CL approach is possibly better than traditional instruction in building up a close association between the learning of English and students' identities. However, it may need more intensive and longer CL intervention to strengthen the link between learners' identities and their learning of English in order to achieve a statistically significant impact.

Introjected motivation stresses the role of internal pressures in motivating learning. The finding that the CL approach was not superior to traditional instruction in increasing learners'

introjected motivation may be accounted for by two factors. First, as with the previous two types of extrinsic motivation, the use of the same learning materials, teaching resources and instructional objectives may account for this result. Second, CL intervention did not act as a catalyst for increasing students' feelings of shame, guilt or embarrassment if they did not know or speak the language well. The primary purpose of CL is to create a supportive and non-stressful learning environment and enable students to improve the target language in an enjoyable and positive manner rather than intensifying those negative feelings about themselves.

External motivation is directly connected with such external pressures as completing a course, achieving a reward, getting a well-paid job, or avoiding a punishment. Team grades, improvement points and bonus points, which are very typical external motivators, were used in this research as important aspects of the adapted CL method, with the aim of generating group cohesion and positive interdependence. The use of external motivators to reward team improvements, successes and cooperative behaviours should have the power to generate more external motivation from students (Jacobs & Goh, 2007; Johnson et al., 1998; Kagan & Kagan, 2009; McCafferty et al., 2006). But the findings of this research did not show this tendency because no difference was found between the two groups in terms of increases in external motivation. The explanation of this result may have been associated with a couple of factors. First, team scores, which were calculated on the basis of team grades, improvement points and bonus points, only made up 30 percent of the final course evaluation, with the major portion of 70 percent allocated to individual scores from the final examination at the end of semester. This proportion of team scores to individual scores in the final evaluation, plus the short-term use of group rewards for only one semester, might not have made group rewards into a major and stable external motivator for the learning of English. In addition, a more immediate explanation was that none of the items on the external motivation subscale of the LLOS were directly related to group rewards or team success. Among the seven items, only two were somewhat relevant, one in relation to academic requirements and the other with course credits. However, these two items were also very similarly relevant to the comparison group since English is learned as a compulsory course for all College English learners. Therefore, it is perhaps understandable that the LLOS-based data did not reveal much difference between the

two groups in terms of change in external motivation in spite of the use of group rewards with the intervention group.

Another point worth discussion here is the appropriate use of group rewards. Some leading CL researchers (e.g. Jacobs, 2006; Kagan & Kagan, 2009) advocate that caution should be exercised when using group rewards because the constant use of external motivators may undermine students' intrinsic motivation. Also, assigning all group members the same grade may cause students to experience feelings of unfairness due to students' differential academic levels. However, this assumption was not supported by the findings of this research because the CL students' intrinsic motivation increased substantially instead of being eroded and also the students were not found to show negative attitudes towards their sharing of work and rewards with teammates.

The positive use of team grades may have been related to several factors in this research. First of all, team grades were not results solely from summative assessment, but were based on formative assessment including feedback from the teacher, peers from other teams and their own teammates. The whole process of grading was actually a combination of formative and summative assessment, with the former as the foundation for the latter. This helped reduce the power of group grades as extrinsic motivators since formative assessment brought students' attention to the reflection on their learning process instead of just focusing on a given grade. It is widely believed that the use of formative assessment is greatly conducive to the development of intrinsic motivation in the long run (Boud et al., 2001; Falsgraf, 2009; Harmer, 2007). Second, group rewards may not do any harm to students' internal motivation if teams are well formed and team tasks are interesting and relevant to students who therefore feel enthusiastic to engage with others. An appropriately designed external motivator can actually serve as a good trigger for learners' intrinsic motivation. In this research, team rewards were found to act as powerful incentives for students to work cooperatively in teams at the initial stage. Over the course of 18 weeks, the inherent nature of CL teamwork of being interesting, enjoyable, stimulating and meaningful shifted students from originally focusing on group rewards to becoming intrinsically motivated to learn. In other words, team rewards initially built in a good reason for students to collaborate and

achieve, but this external motivating power became minor in comparison with inherent enthusiasm generated within the groups when they experienced fun, satisfaction and success with CL and felt like having more of that kind of experience. This change in the students was partly illustrated by the different extent to which they showed care about the accuracy of their team grades and points between the initial and later stages. In the beginning, when the teacher occasionally failed to make an immediate or correct update on team points, the teacher's "carelessness" was quickly spotted by students because they checked the teamwork record sheet frequently with great care. At the later stage, students were seldom found crowded in front of the record sheet posted on the wall or asking the teacher for an addition of one or two points for their teams' previous answers. For a few times at the later stage when the teacher did regular double-checks of teamwork records, she found a few mistakes in her calculations of team grades and points and felt surprised that students did not immediately alert her to those mistakes as before. This change in students' behaviours served as supporting evidence that team grades were less of a focus at the later stage and that students' main interest gradually moved to the process of completing CL tasks along with teammates. In some sense, it is assumed that if the use of group rewards was removed at that stage, it would not change much the students' intrinsic motivation to learn. This finding supported the view that "building in external reasons for students to cooperate can lead to internal motivation to work in groups" (Dishon & O'Leary, 1998, p. 58).

In addition, three specific measures taken in this research were helpful in reducing students' feeling of unfairness about the use of team rewards. First, the use of team rewards was based on careful formation of teams with students' preference for their teammates as well as inter-team homogeneity taken into consideration. When students worked with peers whom they preferred in a team which was of similar academic level to other teams, the feeling of fairness was very likely to be increased. Second, integrating improvement points and bonus points into team rewards created a fair assessment setting that valued students' improvements and efforts rather than their base-line academic levels in comparison with others. This consequently improved students' perceptions of fairness due to their access to equal opportunities for success. Third, an appropriate proportion of team grades in the overall course evaluation was also an important issue. Perceptions of unfairness

cannot be avoided if students' performance on the course is totally or mainly assessed according to team rewards. Generally, team rewards should not take up more than 50 percent of the overall evaluation. Therefore, a comparatively low proportion of 30 percent may have been another contributing factor for students' perceptions of fairness about the use of group rewards in this research.

The last aspect of motivation investigated in this research was amotivation, which stands in contrast to both intrinsic and extrinsic motivation and focuses on a lack of intention and overall motivation (Gagne & Deci, 2005; Noels et al., 2000). Typically learners with strong amotivation do not know why they are learning the language. In this research the use of a CL approach and traditional instruction did not result in differences in students' amotivation. This finding seems to contradict the assumption that CL should decrease amotivation due to its positive effect on students' intrinsic motivation. There were a couple of factors that might account for this result. First, an examination of the mean values of this variable on the pre-test shows that both groups reported similarly low levels of amotivation, that is, an average value of nearly six on a seven-point scale. This finding is not a surprise due to the fact that English is a compulsory course for all Chinese students from elementary to tertiary education. Therefore, with little space allowed for improvement, it is understandable that it was more demanding for students to improve from the point of six to seven than from three to four on a seven-point scale. This phenomenon that participants score at or near the high end of the possible range and have only small opportunities to improve, is referred to as the ceiling effect (Gay et al., 2009). This effect is especially true with Chinese students, who are under the strong influence of the Golden Mean of Confucianism and therefore are most likely to avoid the two extreme poles at either end of the Likert scale. Second, the improvements in intrinsic motivation might not have been influential enough to cause a substantial change in amotivation and overall motivation. This assumption was supported by the finding that the CL approach did not show much advantage over traditional instruction in improving students' overall learning motivation.

Social skills

One major finding on the Social Skills Scale for Chinese College English Learners (SSS-CCEL) is that the CL approach was more effective than traditional instruction in improving students' overall social skills. As for the eight subscales of social skills, the findings can be categorized into three types. First, in the area of equal participation and individual accountability, CL was found to be substantially greater than traditional teaching in improving students' relevant skills. Second, in the three areas of self-confidence, sense of cohesion and checking for understanding, there was a clear trend of greater gains in favour of the CL approach due to the fact that the intervention group improved more than the comparison group according to t-test results and effect sizes. So it is possible that a longer intervention would result in substantial between-group effects in these three areas. However, in the other four areas (i.e. initiative in socialization; being positive; acceptance and empathy; and conflict management), there were no differences between the two approaches in improving relevant social skills. These findings and possible reasons for them are discussed below.

The finding that CL was better than traditional teaching in developing students' overall social skills supports the assertion of many CL advocates that one positive outcome derived from cooperative peer interaction is learners' enhanced interpersonal and social skills (Baloche, 1998; Brown & Thomson, 2000; Gillies, 2007; Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994; McCafferty et al., 2006; Slavin, 1995). Supporting evidence was also provided by a previous study which found that Thai EFL students in the CL classroom demonstrated considerably more positive social behaviours towards second language learning than those in the traditional classroom (Waugh et al., 2005). However, this previous study did not give any details as to what those social behaviours specifically included. Unfortunately, so far there have been very few experimental studies which have investigated the impact of CL on language learners' social skills in comparison to traditional methods.

An important finding from this research is that CL was superior to traditional teaching in cultivating students' skills in equal participation and accountability. This superiority of CL was

linked to its two basic principles—equal participation and individual accountability—which were thoroughly integrated into the adapted CL method used in this research through such techniques as randomly selecting team representatives and assigning individual roles and tasks. This enabled students to know their individual responsibility for teamwork, understand the importance of their personal contribution to team success, and thus feel obliged and motivated to participate actively and equally. Individual accountability greatly facilitated equal and active participation among students, and equal and active participation contributed to successful teamwork and learning. This may serve as another important explanatory factor for the finding that the intervention group achieved greater gains in English learning than the comparison group.

The CL approach was found to be better than traditional methods in improving students' self-confidence in this research, although this difference did not reach statistical significance. This finding supports widely accepted views among CL advocates that cooperative experiences result in greater psychological health and higher self-esteem, of which self-confidence is a major component (Gillies, 2007; Jacobs & Goh, 2007; Johnson et al., 1998; Kagan, 1994; Kagan & Kagan, 2009; Slavin, 1995). However, this is somewhat inconsistent with the findings from Ghaith's (2003) ten-week comparative study regarding teaching reading in English to Lebanese high-school students, which found that there was little difference between CL and traditional methods in developing students' academic self-esteem. Ghaith suggested that a longer intervention may be necessary for greater gains in the area of students' psychosocial adjustment. The current research involved an 18-week CL intervention, which nearly doubled Ghaith's intervention length and was also integrated into a wide range of instructional areas: listening, speaking, reading, writing and vocabulary. Consequently, this might have resulted in a more powerful intervention than that in Ghaith's study and thus generated a more positive outcome. However, it is worth noting that although there was a clear difference between the two methods on the variable of self-confidence in this research, this difference was not statistically significant. A statistically significant between-group effect may call for a longer intervention, which again agrees with Ghaith's view that an adequate length of CL intervention is important to produce reliable results on some variables especially those relating to some psychosocial factors. In addition, there is an

assumption that CL is likely to promote students' sense of accomplishment and self-worth, which make up an important contributing element of self-confidence (Holt, 1993; McCafferty et al., 2006). Thus, the intervention group's improvements in their English learning might also partly explain their enhanced self-confidence in this research.

In this research, the CL approach produced more positive effects than traditional methods on students' sense of cohesion, although this difference did not reach statistical significance. This finding is consistent with the proposition that positive interdependence, a fundamental principle of CL, is greatly conducive to building up supportive peer relationships and group cohesiveness, and thus facilitates students' sense of cohesion. This result regarding a sense of cohesion is somewhat inconsistent with Ghaith's (2003) study, which found little differences between CL and traditional teaching in affecting Lebanese English learners' feelings of school alienation—a variable closely related to students' sense of cohesion because feelings of school alienation undoubtedly undermine a sense of belonging. As for the cause of this disagreement between these two studies, the explanation made above about conflicting results on the variable of self-confidence between the two studies may be similarly applied to this situation again; that is, this current research involved longer and more intensive intervention than Ghaith's study. Also, as with the variable of self-confidence, a statistically significant between-group effect on the variable of sense of cohesion may require a long-term intervention with CL.

Checking for understanding is fundamental for meaningful interaction and successful communication. In CL teamwork, students feel motivated and obliged to use skills such as offering or asking for explanation, clarification, elaboration and examples, with the aim of making oneself understood by teammates as well as understanding others, and further achieving team success. This may explain the finding from this research that the intervention group improved relevant skills better than the comparison group, although this difference did not reach statistical significance. In addition, the heterogeneous nature of CL teams also enhanced the necessity for checking for understanding among teammates. In contrast to the situation that checking for understanding was used at every stage of teamwork in the CL classroom, peer interaction seldom occurred in the traditional classroom and checking for understanding was mainly used top-down

from the teacher to students, which is typically characterized by a routine question like “Do you have any problem in understanding what has been taught?”. As with variables of self-confidence and sense of cohesion, a statistically significant between-group effect on the relevant skills in checking for understanding may need the involvement of a longer or more intensive intervention.

The findings of this research did not show differences between the two methods in cultivating students’ skills in the other four areas: initiative in socialization; being positive; acceptance and empathy; and conflict management. These findings do not support the assumptions of some CL advocates that the use of cooperative interaction benefits the development of relevant skills in those areas (Baloche, 1998; Brown & Thomson, 2000; Gillies, 2007; Johnson et al., 1998; Kagan, 1994; Slavin, 1995). But so far, in spite of these assumptions, no experimental study other than this current one has been found which compares the effects of CL and traditional teaching on these relevant variables. In this research, a couple of factors might account for the findings concerning these four areas. First, these four areas are all related to students’ personality and disposition, which are located in their nature and thus hard to change through short-term interventions. So it is understandable that an 18-week CL intervention was inadequate to make a marked difference in these areas. Second, the lack of sensitive items might compose another explanatory factor for these findings. In this research, the use of CL interaction was restricted to only regular teammates in the teaching of College English course, but only two out of the 20 items in these four areas were directly related to English teaching and within-team cooperation. For instance, students were likely to give more positive responses when an item like “I offer help to those who cannot grasp materials learned” would be changed to “I offer help to teammates who cannot grasp materials learning in the English course”.

In addition, it should be noted that the development of social skills in this research was treated as a by-product and a natural consequence of the implementation of CL because, during the whole process of the empirical study, there was no involvement of any particular instruction specially for the purpose of teaching students social skills. So considering the economical nature of this intervention in terms of enhancing social skills, the overall outcomes on the SSS-CCEL could be considered quite positive and encouraging.

English-Learning-Related Social Skills

On the Social Skills Scale for Chinese College English Learners (SSS-CCEL), there are nine items which are particularly related to the learning of English and thus are considered more sensitive to the effects of the intervention on students. They are:

Item 3: I am doing a good job of English language learning.

Item 4: I feel I am making progress in English language learning.

Item 7: I consider peer support indispensable to my English language learning and success.

Item 10: I offer teammates support and assistance as much as I can, so that they do their best in English.

Item 15: I feel stressed and uncomfortable when working with others in English classes.

Item 25: I often use examples to make myself understood in English classes.

Item 26: I have a clear picture of my personal role in English teamwork and participate actively.

Item 27: I tend to keep silent in English classes, except when I am called to answer questions.

Item 35: I believe everybody of different English levels can make a contribution to the completion of group tasks.

The findings of this study based on these nine items generally indicate the superiority of CL over traditional teaching in developing students' relevant skills because the intervention group improved more than the comparison group. According to the different extent to which CL was better than traditional teaching in enhancing those skills, findings can be classified into two categories. The first category involves items 4, 10, 15, 26, 27 and 35, where CL was found to be substantially better than traditional teaching in improving relevant skills. The second involves items 3, 7 and 25, where the use of the two methods did not result in substantial differences.

The data based on item 4 indicate that the intervention group perceived that they had made more progress in English than the comparison group. One's perception of making progress is a very importance source of one's sense of achievement, which greatly contributes to the development of self-confidence (Johnson et al., 1998; Kagan, 1994; Sharan, 1994). This is

because, when students feel or think they are making progress, they are gaining a sense of self-worth and achievement, and thus feel more self-confident. In turn, this enhanced self-confidence may generate more progress and a stronger sense of achievement. These assumptions are supported by the findings of this research that the intervention group made substantially greater improvements than the comparison group in their academic achievements in English, as well as in their perceptions of making progress and their sense of self-confidence.

The findings regarding item 10 and 15 that the students in the CL classroom were much more likely to support teammates' learning and feel comfortable when working in teams, are in accordance with the widely accepted assumption that CL facilitates a more supportive and safe environment where peers are positively interdependent and thus care about each other as well as each other's learning (Johnson et al., 1998; Kagan, 1994; Kagan & Kagan, 2009; McCafferty et al., 2006; Slavin, 1995). These findings also agree with a previous study, which compared the effect of CL and that of whole-class instruction with secondary-school EFL learners in Hong Kong, and found the majority of students in the CL classroom were positive about the use of CL since it created a favourable learning environment of being supportive and non-threatening (Sachs et al., 2003).

Another interesting finding regarding item 15 is that considerably more stress and discomfort were developed within the comparison group following the intervention. This finding is actually quite a surprise due to the expectation that the traditional group work used with the comparison group should be of less stress to participants because they were allowed to work with neighbours they preferred and then volunteer to report back instead of being randomly selected to represent teams as was used with the intervention group. However, a close study of data on this variable found that both groups reported a high level of comfort and ease with group work at pre-test, especially the comparison group which registered a mean value approaching six on a seven-point scale. This result is unlikely to reflect the truth of their experience due to the well-known fact that teacher-talk mostly dominates EFL classes (Bailey, 2003; Harmer, 1998, 2007; Nunan, 2003a) and group work is very rarely used in the EFL classroom in China (Chen, 2007; Hu, 2002; Hu, 2005; Jin & Cortazzi, 2004; Shu, 2004; Siemon, 2010; Teng et al., 2004; Wang, 1999). Students' comfort

and ease with group work calls for adequate time and experience to get accustomed to it, which is obviously unavailable to Chinese EFL learners in most cases. So the high mean values on this variable may serve as a reflection on students' expectation of their performance in group work rather than what actually happened. This could also be an indication that Chinese students felt positive about group work and wanted to have more experiences with it (Zhang, 2006). However, when those in the comparison group were face to face with unstructured traditional group work, their feeling and experiences might appear rather different from what they had expected. It might not only involve some fun, but also a lot of negative things; for example, high-achievers dominated groups with little consideration for others who might consequently feel inadequate and then lose interest in participating and learning in groups. On many other occasions, when group members did not have the necessary competence in self-regulation and learning autonomy—which is a common problem with Chinese students—students might quickly get off task or fall silent for lack of a clear-cut procedure or structure to follow. The teacher did notice that some students in the comparison group went back to work individually on topics or tasks, which clearly registered their negative attitude to group work and its effectiveness. All these factors may account for the rise in the comparison group's discomfort and unease with group work. On the contrary, the intervention group used with CL teamwork showed a very different trend on this variable which suggested a substantial improvement in students' comfort with teamwork at post-test, in spite of a similar overstatement on their comfort with group work at pre-test. These contrastive findings between the two groups actually provide evidence for the superiority of CL teamwork over traditional group work in teaching English to Chinese students who are new to group work. In addition, students' overstatements on this variable at pre-test reveal the weakness in the use of self-reported questionnaires, which may produce inaccurate or misleading data. For instance, in this case, students may have mixed up what they expected of themselves with what they actually could do. This constitutes one of the limitations of this research, which will be further dealt with in the section on limitations of this research.

The results from item 26 and 27 suggest that the intervention group substantially outperformed the comparison group in the understanding of their personal responsibility for teamwork,

engagement of active participation and quantity of student talk during the class. These findings directly represent positive consequences of three basic CL principles—individual accountability, promotive simultaneous interaction, and equal participation—to students’ learning. They support the common assumption that CL activities may greatly enhance student talk time (STT) and active participation on the whole (High, 1993; Holt, 1993; Kagan & Kagan, 2009; McCafferty et al., 2006). In second or foreign language learning, adequate amount of STT is considered a critical factor in improving target language proficiency, especially in speaking, and thus “the best lessons are ones where STT is maximised” (Harmer, 1998, p. 4). In addition, these findings also agree with a previous study which compared the use of three teaching modes—direct instruction, unstructured group work and structured CL group work—on student participation in the FL classroom, and reported the superiority of structured CL group work over the other two in generating output and active participation from students (Magee & Jacobs, 2001). In this current research, apart from the use of structured CL teamwork, another factor that might have facilitated the increase of the intervention group’s oral output production was the use of bonus points in the course evaluation. This technique provided team members with an external reason to encourage each other to volunteer answers to the teacher’s questions, and resulted in more voices and thereby active participation in the CL classroom.

According to item 35, the intervention group substantially strengthened their belief in the contribution each team member could make to team success in spite of their different English levels. This partly reveals that low-achievers, like other teammates, also found their positions and roles in CL teamwork, and that their effort and work were valued as an integral part of team success. This formed a contrast with traditional group work which is usually dominated by high-achievers (Pica & Doughty, 1985), while low-achievers tend to be neglected or avoided and they thus usually feel demotivated or stressed when working in groups. This finding supports the view of Sharan and Shachar (1988) and Cohen et al. (2004) that CL creates a more congenial and accepting language learning atmosphere where students at the disadvantaged end of the spectrum, either in terms of academic or language level, can make fuller use of their language abilities and gain more positive learning experiences than in traditional settings.

The finding regarding item 3 suggests that no substantial differences were found between the two groups in their responses to the question “I am doing a good job of English language learning”. In other words, the use of two different approaches did not result in much difference in improving students’ satisfaction about their English learning. This seemingly contradicts the findings that the intervention group perceived they were making greater progress in English learning (item 4) and indeed achieved significantly more than the comparison group according to the results on the CET. However, this seemingly contradictory finding is perhaps understandable if several factors are taken into consideration. First, while students are making progress, they are likely to set up new learning targets at a higher level. Consequently, what they have attained always tends to be inadequate in their eyes if they aim higher and have the potential to achieve more. Second, as a result of over ten years of exposure to competitive educational contexts, students were used to associating the concept of “doing a good job of English language learning” more with their learning results in comparison with peers, rather than linking it with their own progress. It is understandable that a short-term intervention can hardly make a change either to this long-standing perception of students or to student rank-order in the subject of English. Third, results on the CET showed both groups were generally at a similar level of English proficiency since their overall mean scores were only different by five points out of 100. This means, in general, the overall quality of English learning was comparable between the two groups. So it is understandable that the two groups’ general comments on their learning quality were not much different. Last, this finding might be partly derived from a factor relating to Chinese traditional values and ways of doing things. Modesty is traditionally valued in the Chinese culture, which sometimes goes so far that people do not easily admit their own success or excellence. A simple example is that Chinese students typically respond to others’ praise or compliments like “You are really good at maths” with “I’m just OK, not good enough” instead of accepting it proudly with a “Thank you”, even if they do indeed excel in maths. This phenomenon, usually termed as “self-deprecation” or “excessive modesty” may result in students’ habitually downgrading their learning. In this case, this habitual self-deprecation could have more consequences for those students who actually did a good job, because, in spite of this fact, they were still likely to give

similar responses to those provided by students not doing a job.

The two groups' responses to item 7 regarding the importance of peer support in their English learning and success did not demonstrate substantial differences. This somewhat conflicts with the assumption that promotive peer interaction inherent in CL may increase students' awareness of the vital role of peer support in their language learning, especially in the FL learning context (Hammer, 1998, 2007; High, 1994; Jacobs & Goh, 2007; Kagan & Kagan, 2009; McCafferty et al., 2006; McGroarty, 1993). One possible explanation for this seemingly contradictory finding could be related to the fact that this short-term CL intervention only involved peer interaction within teams of foursomes; that is, students only had access to support and facilitation from three teammates out of around 50 classmates. This weak version of peer support within foursomes might not have been powerful enough to generate a striking change to students' perspectives on its importance. In the same vein, this short-term intervention focusing on cooperation within teams in the English classroom may not have been sufficient to demonstrate the potential power of peer support on success as a whole, especially considering that Chinese students are constantly faced with fierce competition in education, which particularly stresses the role of hard work and diligence rather than peer support (Jin & Cortazzi, 2004).

The two groups were not found to be much different in their responses to item 25 "I often use examples to make myself understood in English classes". At both pre- and post-test, both groups registered relatively high mean values, above four on a seven-point scale, which means students in both groups were positive about their skill in using examples to make themselves understood. Using examples to illustrate things or ideas is one basic skill for checking for understanding (Brown & Thomson, 2000; Jacobs & Goh, 2007; Jolliffe, 2007; Johnson et al., 1998; Kagan, 1994). This skill is simple and easy to use in comparison to other complex skills for checking for understanding, such as elaborating and summarizing. This may partly explain why both groups reported frequent use of this skill in English classes. However, since the two groups demonstrated substantial differences in the quantity of STT, output production and active participation, which all involved the "push and pull" for using examples for explanations, there may be a chance of greater gains in favour of the intervention group. In other words, a longer intervention may lead to more

positive results on this variable in favour of the CL approach.

Implications for Practice

Using a new teaching approach is always full of challenges, and this current research is a case in point. The application of CL in the Chinese EFL teaching context turned out to entail various challenges, which included design of appropriate CL tasks, extra workloads involved in CL lesson preparation, limited teaching hours and large amounts of designated teaching content, as well as students' overuse of Chinese in teamwork.

Design of Appropriate Cooperative Learning Tasks

Appropriate learning tasks play a critical role in realizing designated instructional objectives and learning goals. The importance of tasks in the CL classroom cannot be overemphasised, because a CL lesson is composed of a series of CL activities aimed at completing particular CL tasks. Many leading researchers (e.g. Jacobs and Goh, 2007; McCafferty et al., 2006) assume that the failure of CL activities is mainly due to the inappropriateness of team tasks either in terms of students' language levels and interests or inadequacy in the inclusion of fundamental CL elements (e.g. positive interdependence and individual accountability) in the process of completing and assessing the task. Bejarano (1987) maintained that appropriate CL tasks facilitate the processing of both instructional content and social skills in small groups. Therefore, designing appropriate tasks is the key to successful learning and teaching in the CL classroom, and has been well recognized as an essential job for a teacher to undertake.

In the current research, English teaching was based on the assigned textbooks designed mainly for the traditional classroom rather than the CL curriculum, so the ready-to-use activities provided in the textbook hardly suited the CL purpose and thus designing CL tasks became a must-do for nearly all lessons. This is a very common problem not only in China but across many other

countries where traditional whole-class teaching still dominates. In this research, although the instructor of the English course had become familiarized with key CL techniques and obtained some experience with CL classroom instruction through workshops and the pilot study, she still found the job of designing tasks full of challenges. These challenges included choosing a topic which could well meet students' interests and also simultaneously fulfil the instructional aim of the unit, establishing teamwork grading criteria for students to refer to in the process of completing team tasks, and designing worksheets which provided adequate information and scaffolding to guide and facilitate the processing of group activities. Thus, it can be imagined that this job would be very challenging for novice teachers and those who do not have much knowledge and experience in this field. Actually this situation as regards teachers' lack of competence in CL is already partly reflected in recent studies in China, which show that Chinese English teachers and researchers tend to mistake group work for CL and thus are far from being competent in this field (J. Li, 2007; R. Li, 2007; Luo, 2007; Wang, 2004; Zhang & Zhao, 2004).

In view of this issue, some countermeasures are recommended regarding how to support teachers in designing appropriate CL tasks. First, some relevant training programmes and workshops should be organized for teachers who are interested in employing CL techniques in the classroom. Training content should include educational philosophies and basic principles of the CL approach, but it would be preferable to integrate these theoretical conceptions into the actual application of CL tasks. Trainers or presenters should provide practical techniques, demonstrate detailed lesson plans, and ensure that trainee teachers have opportunities to practise the CL techniques they have just learned in the programme. Kagan's (1994) Structural Approach is recommended initially in the programme because of its special feature of providing numerous simple CL structures. These structures are easily used or adapted for constructing CL tasks when learning content is specified, and are functionally flexible for use as part of any lesson. In other words, Kagan's Structural Approach focuses on simple structures that can be used as part of any lesson instead of designing whole CL lessons, so it is more feasible for beginners to start with. As for the appropriate method used in the training programme, it is hard to imagine how a teacher who has been taught CL through lengthy lectures with no involvement of cooperative interaction

can grasp the true essence of CL and use CL tasks effectively in his/her classes later. Thus, it is important that trainee teachers learn about CL by personally taking part in and completing CL tasks with peers, so that they can have positive experiences with CL and improve their confidence in and enthusiasm for using CL tasks with their students (Gwyn-Paquette & Tochon, 2002; Hornby, 2009; Jacobs, Lee & Ball, 1996). However, it must be admitted that these programmes and workshops are not easy to get access to, because of a lack of relevant expertise and a shortage of financial support, which is true for the College English teaching context. It is hoped that this situation can be improved with more professionals and educational leaders realizing the positive outcomes that the CL approach can bring to EFL teaching.

Second, it is strongly recommended that English teachers, who have an interest in CL and use the same sets of textbooks, build up a cooperative teaching team (ideally composed of three or four staff) to support each other in experimenting with CL. They should work together designing tasks, planning lessons, observing each other using tasks in class, sharing good ideas, and helping each other sort out implementation problems. This collective effort and team support among colleagues will expand sources of information and learning, make the experience more positive and productive, and thereby provide an effective means to help teachers better meet the challenge of designing appropriate CL tasks and lessons.

The third recommendation is for those teachers who want to give CL a try but to whom the training programmes, workshops and cooperative teaching teams are not accessible. This is a very difficult situation for them, but it does not mean that the door to using CL tasks is closed to them. Above all, they need some reliable resources for a fundamental knowledge of the CL approach and basic principles for task design. Some classic CL books, such as those written by Kagan (1994) and Slavin (1995) provide practical guides to CL methods and relevant principles. In addition, there are some websites on CL hosted either by leading CL experts (e.g. www.georgejacobs.net/cooperative.htm by George Jacobs) or some CL research centres and organizations (e.g. www.iasce.net by the International Association for the Study of Cooperation in Education, www.co-operation.org by Cooperative learning Centre at the University of Minnesota, www.kaganonline.com by Kagan Cooperative Learning). These websites provide answers for

frequently asked questions relating to task design and lesson plans in the CL classroom, and also offer professional updates and recent research findings, as well as serving as forums for teachers to get support, share and exchange ideas. Additionally, the discussion board is often available online at these CL websites, and CL experts are always there to provide help and support. For instance, a teacher can receive responses within 24 hours from Kagan's online discussion board if s/he posts a question or query there. Thus these websites are helpful resources for teachers to rely on. Another point that a beginner teacher should bear in mind is to start with simple and short tasks (e.g. using Kagan's structures), and choose topics and structures that keep both them and their students within their comfort zones.

Extra Workload Involved in Preparing and Implementing Cooperative Learning

Lessons

This study has shown that that well-designed CL lessons can provide useful learning experiences for students and generate positive outcomes as regards academic achievement, intrinsic motivation as well as the development of social skills. The resultant strong sense of achievement on the part of teachers may make them consider that their hard work and time are well rewarded. However, on the other hand, there is evidence that teachers may "shy away from using cooperative learning activities as they are perceived to be too time-consuming in lesson preparation" (DaSilva Iddings, 2006, p.179). Teachers often find that CL lessons, in comparison with the traditional direct instruction they are used to, involve a lot of extra effort and time. They need to decide on appropriate topics, set up operational procedures, establish assessment criteria, design worksheets, work out a list of scaffolding words and expressions for teamwork, and meet unexpected challenges cropping up in the processing of teamwork.

In view of this, a number of recommendations are provided as potential solutions to reduce teachers' workloads. First, it is helpful if there is a teachers' resource book available to CL teachers. The resource book would demonstrate model lesson plans for each unit of teaching

materials, including detailed operational procedures, appropriate team tasks for particular learning objectives as well as matched facilitative worksheets. This book may serve as a great convenience for teachers to refer to when preparing their CL lessons. The production of a CL resource book is not only helpful but feasible if teaching is based on designated textbooks which may stay unchanged for some years. College English teaching is a case in point, for which there are stable textbooks and national curriculum requirements. So it is feasible that some CL experts work out a set of teachers' resource books based on particular textbooks which teachers can refer to in their classroom practice.

The role of CL teaching teams is also emphasized in reducing teachers' workloads and sustaining teachers' effort in the use of CL (Cohen et al., 2004; Fullan, Bennett & Rolheiser-Bennett, 1990; Jacobs, 2006). There are a couple of ways that teachers can collaborate with each other in teams to reduce workload. The simplest way is to allocate jobs of producing lesson plans to individual teachers in advance. The allocation can be made in accordance with the unit of teaching materials (e.g. each teacher is accountable for one or two units) or teachers' strengths and interests in particular linguistic aspects. For example, a teacher who is strong or interested in teaching speaking is responsible for preparing tasks with a focus on students' speaking competence. This actually splits a heavy workload into several portions and thus effectively relieves teachers of excessive work. It should be noted that an allocation of jobs does not mean individuals working alone, as it is essential that channels for discussion, consultation and support are always open within the team. In a similar vein, the use of team teaching, with each teacher responsible for the design and delivery of part of a lesson serves as another possible means of reducing teachers' workload. Team teaching is also a very powerful way for teachers to model cooperation in front of their students through demonstrations of how to contribute as part of the team by collaboratively planning, reflecting, communicating, making decisions and reaching a compromise (McCafferty et al., 2006; Murdoch & Wilson, 2008). Students always learn more from what teachers do than what they say. However, it is also realized that, in such educational settings as College English teaching, the structure of team teaching is often a luxury beyond reach because of shortages of English teachers.

When the above two types of support—a reliable teacher resource book and CL teaching teams—are not available, teachers with an interest in using CL in their teaching would have no choice but rely on themselves. In this case, they should not expect to throw out their traditional lessons and start with the design of complex CL lessons, for the resultant frustration and stress may lead to a quick waning of their initial enthusiasm and even giving up the use of CL completely. Instead, it is strongly recommended that they include CL teaching as a small part of lessons only on an occasional basis, for instance, using a five-minute task organized using Kagan's structure of Think-Pair-Share once every session. A proportion of five minutes of CL versus 45 minutes of traditional teaching does not form a radical change to the classroom setting, and is unlikely to lead to discomfort or uneasiness of both teachers and students. As teachers and their students become more comfortable with the simple CL techniques, they will feel motivated and confident in the use of more complex techniques on a more frequent basis. A gradual increase of CL teaching within both teachers' and students' comfort zones is essential for a successful transit from the traditional to CL classroom setting. Inexperienced teachers may need more time for this transfer. However, the positive thing is that no matter how slow this process is, they have started the exciting explorative journey to include an innovative approach into their own teaching repertoires.

It would be ideal if schools or universities show their support and appreciation for the teachers' extra effort and time spent on CL lessons. This can be expressed by some particular means like awarding them special annual bonus payments, presenting certificates of recognition, allowing them the privilege of attending academic conferences, participating in professional training programmes and getting access to professional promotion. Undoubtedly, the fulfilment of this recommendation involves a range of issues relating to administrative procedures, current financial budgets, and especially all sorts of institutional and organizational support. These issues tend to be rather complex and are often beyond the control of individual teachers and also beyond the scope of this thesis.

Additionally, as for a FL teacher, their workloads are likely to increase owing to some unexpected linguistic challenges occurring in the process of facilitating teamwork, where the

teacher acts as a language facilitator to provide on-site scaffolding. The following is a possible scenario in the CL classroom, where students are completing a team task on the topic of their favourite sports stars. Students who choose some foreign players (e.g. Kobe Bryant who plays shooting guard in the NBA for the Los Angeles Lakers) may have problems in saying correctly in English the players' names, positions and host teams, and then they turn to the teacher for help. Unfortunately, the teacher may not be able to give satisfactory answers either. Actually, no matter how hard a teacher has worked to prepare a comprehensive list of scaffolding words and expressions, it may not include all students' interests. Traditionally in China, teachers are considered as gurus or sages of particular subjects, so they are very likely to feel embarrassed for failing to provide answers and fear losing face in front of their students. This may constitute a reason why so many teachers constantly enjoy giving prepared lectures even if they are well aware of its inadequacy and ineffectiveness in teaching a language. In other words, by simply delivering presentations on the textbook materials, teachers actually keep a safe teaching environment for themselves because they do not interact with students and thus avoid these unexpected challenges which often occur in the CL classroom. CL teaching aims to create a comfortable and safe learning environment for students, but it does not mean that CL teachers have to feel nervous and stressed about facing unexpected challenges from students. It is strongly recommended that, for the benefit of both teachers and students, these challenges should be taken positively as a chance for the teacher to set up a good role model by demonstrating how to deal with challenges properly. First, they should gracefully admit to their students that they are not a sage and they do not know everything or get everything correct. This actually helps teachers step down from the sage's platform and become more equal with and close to their students, which may generate more interaction between the teacher and students, help students face their failures and errors properly, and further motivate them to take risks and try out the language. All these are just what is desired and greatly valued in the contexts of both CL and language teaching. Second, teachers' admission of their inadequacy is not sufficient, and more importantly, effort should be engaged to work out solutions as soon as possible. Teachers could initiate an immediate five-minute query session following teamwork, in which teachers present to the whole class remaining problems and ask

questions to invite students' ideas and contribution. If the problems are successfully solved through some students volunteering their answers and ideas, their host teams obtain bonus points. More importantly, this illustrates the power of collective effort through a good example, and serves as convincing evidence that there is always something that one can learn from others because even the teacher is learning from their students. If the problems cannot be solved in class, teachers should get access to all possible resources (e.g. dictionaries, colleagues, native-speaker friends, on-line resources) after class and bring the answers to students for the next time. As a matter of fact, these challenges do provide teachers with more opportunities for learning and broadening their knowledge.

Lastly, teachers' workloads can be considerably reduced as both teachers and students gradually gain their skills and experiences in CL (Cohen et al., 2004; DaSilva Iddings, 2006; Jacobs, 2006; Joritz-Nakagawa, 2006; Wee & Jacobs, 2006). For instance, the carefully-prepared worksheets were sometimes replaced with a simple team-grade sheet (see Appendix 5) at the later stage of the current research when students were used to the teamwork procedures and routine assessment criteria. Many CL practitioners have the experience that "once the necessary infrastructures for cooperative practices are in place within a classroom community, it is relatively simple to implement lessons" (DaSilva Iddings, 2006, p. 179). Along similar lines, Joritz-Nakagawa (2006) depicted her exciting and successful progression with CL teaching as follows:

I used to use cooperative learning sparingly, but now it is my primary approach. This progression was the result of gaining confidence and experience, as I continued to experiment with adding cooperative learning activities to my courses, encouraged by the positive reactions of students and of colleagues with whom I discussed my courses ... This is not to say I walk into the classroom with a lockstep notion of what to do. These days I walk in with more of a sketch and adapt it to what I see occurring in the classroom. I know now that the best class period I teach is rarely the result of painstaking advance planning, but more likely the result of minimal planning and my being able to adapt on the spot to what I see occurring, whether positive or negative, in the class. The difference now is that I

have many trusted cooperative learning tools to pull out of my hat, as needed. These are tools I have used so often that now they are second nature. When new to teaching, I often made detailed lesson plans, writing out each step of my lesson in the order I thought it should best proceed. This is now rarely the case. In other words, I have reached what could be called unconscious competence in cooperative learning (p.136)

Actually, the English instructor and researcher in the current research had a similar experience that lesson plans became easier and less time-consuming with the passing of time. She also found herself to be more and more comfortable and confident in dealing with all kinds of challenges in the process of facilitating team tasks, changing from feeling nervous when students raised challenging questions to feeling pleased and proud of students' active learning and critical thinking, which is actually the true value of education. This positive and comfortable feeling which the teacher developed when working with students in the CL classroom made the job enjoyable, reduced work pressure, and thus lightened the workload in some sense. So it is believed that the problem of heavy workload is likely to be overcome with time, and also the fundamental issue of decreasing the workload lies in the teachers' perseverance in integrating CL techniques into their own teaching repertoires. In a teaching context like College English teaching, where the same textbooks are used for quite a few years, teachers would find the job a lot easier in their second or third time of teaching the same material, for quite a lot of tasks can be reused or slightly modified.

Limited Teaching Hours and a Large Curriculum to Cover

Teachers often feel a lot of pressure to cover the teaching materials within limited teaching hours, especially when there are required curriculum and textbooks to follow. Thus teachers may drop CL soon after having several trials, for they find teamwork takes time and is not as quick as teacher-fronted lectures for covering the curriculum. However, one important point to bear in mind is that using CL does not mean abandoning whole-class direct instruction. Many CL experts suggest adding CL tasks gradually in the traditional classroom (Jacobs & Goh, 2007; Jacobs et al.,

2002; Kagan, 1994; Kagan & Kagan, 2009; McCafferty et al., 2006). Initially, teachers should try CL in a very limited way, for example, using a simple CL technique for three to five minutes in a 50-minute session, and this has little impact on the speed of covering curriculum. According to Kagan (1994, p. 1:4), after teachers have gained a good command of “the art of managing a classroom of teams and feel competent in one structure, you (teachers) may well begin to include other techniques—eventually finding the amount and style of grouping which best fits your own style”.

The inefficiency of CL in delivering teaching materials is partly due to the factor that teachers and students are not competent and efficient in dealing with teamwork, and sometimes feel lost at the initial stage when they are not used to the CL techniques and procedures. One encouraging situation is that, with students’ improvements in their familiarity with teammates, experience in CL procedures, as well as relevant cooperative skills, they are “more adept at figuring out what needs to take place for groups to swing into action” (Jacobs, 2006, p. 44). Thus the time spent on the completion and processing of tasks can be greatly reduced at the later stage. This situation is well reflected in this current research, where teams were found functioning much better and quicker in the second half of the semester and thus more and more teamwork was involved in a lesson.

In L2 teaching, traditional whole-class instruction has strengths that CL teamwork does not provide. For instance, it serves as an efficient way for teachers to explain linguistic knowledge and culture points, including the use of a new word, a complex grammatical point, or relevant background information that cannot be obtained through student peer interaction. Some CL advocates (e.g. Murdoch & Wilson, 2008, p. 5) propose that the CL approach is not “a total classroom program or organization strategy” and it “works best if combined with other approaches”. This combination is well reflected in a major CL method—STAD, which includes the traditional class presentation by the teacher as one of its five essential components. In addition, this present research, using STAD as the basic format for its CL adaptation, also provides further evidence that whole-class direct instruction can form an integral part of a well-designed CL method. Actually, these two strategies—CL and whole-class teaching—are more accommodating

than conflicting. The proportion of whole-class instruction to CL can be flexibly adjusted in line with the objectives of lessons, teachers' expertise at preparing and implementing CL lessons, as well as students' familiarity and comfort with teamwork.

In addition to the appropriate combination of CL with whole-class teaching, a number of techniques are also particularly recommended for the purpose of improving the efficiency of teamwork. First, time limits should be set for tasks, so that students can budget their time accordingly. Although there is evidence that specifying a time limit may inhibit complexity and accuracy of learners' output production (Ellis, 2009c), giving students an unlimited amount of time to perform a task is not feasible in classroom settings due to limited teaching hours. Thus, teachers should fully use their common sense to set reasonable time limits. Second, each team should delegate one member to be the time monitor or taskmaster, whose responsibilities are to insure the team stay on task and get the job done within time limits. Choosing a team captain is also very helpful in improving teamwork efficiency since captains assist the teacher to supervise and monitor the functioning of their own teams and keep teamwork on the right track. Third, tasks should not be too difficult, either cognitively or linguistically, so that students can handle them through cooperation, while difficult learning materials should be explained through teachers' whole-class presentation. Ideally, all the information regarding time limits, relevant roles, task requirements, scaffolding phrases and assessment criteria should be included in a well-designed worksheet to be distributed to each team for the convenience of their teamwork. Fourth, bonus points for teams can be used as a technique to stimulate teamwork efficiency. In some cases, the teacher can allow a team that first completes the task to volunteer answers and win special bonus points for the team. This provides an incentive for team members to support each other and work efficiently.

Lastly, there is one critical point that should not be ignored, that is, the ultimate purpose of L2/FL teaching is not merely to cover curriculum but to develop students' communicative competence. There is clear evidence that teaching English in the traditional approach year after year does not effectively improve students' communicative competence but largely contributes to rote learning. CL teamwork is considered an essential technique for L2/FL teaching because of its

potential for fulfilling the ultimate teaching goals. This is in spite of the possibility that the use of teamwork slows down the processing of learning materials in some cases and cannot cover as much curriculum as direct instruction does.

Students' Use of the First Language in Teamwork

Students' use of their first language in teamwork should be treated as a normal phenomenon in the second language classroom (Ellis, 1999). A limited use of the first language is conducive to relieving students' excessive anxiety and stress (Burden, 2000; Harmer, 1998, 2007; Jacobs & Goh, 2007; Jacobs, et al., 2002; McCafferty et al., 2006). But the problem is that students may overuse their first language to complete the task quickly. In this case, some measures must be taken to alleviate the problem.

First of all, teachers should make it clear at the beginning of a course that the main purpose of team tasks is to offer students chances to practise English with peers in authentic situations, rather than completing tasks for the sake of it. It is stressed that "using the first language as a shortcut to completing the task actually defeats the main purpose of the task" (Jacobs & Goh, 2007, p. 44).

Second, when the overuse of their first language occurs with a good number of students, some reflection should be made by teachers, for instance, on whether the task is too difficult or whether necessary scaffolding phrases are not available to students. It is also very important that teachers set up good examples to maximize the use of English either when interacting with students or delivering lectures.

Third, when students feel like using their first language, they are told to write it down for team members instead of saying it. This strategy involves three benefits: writing does not disturb other teams or affect the learning atmosphere of using English as the main communicative tool; writing takes time and thus puts a curb on students' overuse of their first language; and importantly, it provides a convenience for teachers to observe teams' performance, that is, whether they are busy writing or speaking to each other.

Lastly, a rule should be set in the very beginning that the final team product must be presented

in English, which guarantees that team members use a considerable amount of English in processing their team tasks. Meanwhile, a reward incentive should be included in the assessment system to stimulate the use of English. For example, consistent use of English in teamwork should be recognized with bonus points.

Contributions of this Research

This research, involving a 13-week pilot study and an 18-week main study, has made some contributions to the evidence-base in the area of CL and language teaching. First of all, the main findings of the research supported the effectiveness of CL in teaching English to tertiary learners in China. The findings were derived from a comprehensive investigation of a wide range of areas on three measures: the College English Test (CET), the Language Learning Orientations Scale (LLOS), and the Social Skills Scale for Chinese College English Learners (SSS-CCEL). The general conclusion of this research is that CL may have an important role to play in College English teaching in China.

The research also offered a detailed model format for CL adaptation in accordance to the EFL teaching context, which increased the replicability and reliability of this research. This model includes team formation, technique adaptation and course evaluation, which turned out to suit the College English teaching carried out in this research. This is in spite of challenges such as the design of textbook-based team tasks, large-class instruction, tight teaching syllabi and limited teaching time, as well as students' unfamiliarity with CL skills and limited learner autonomy, which are considered common features of EFL teaching contexts in many Asian countries. So this adapted CL model may be of considerable interest to colleagues who work in similar educational contexts and are seeking innovations in order to improve educational outcomes. The successful use of this model in this research may provide insights for colleagues into their future experiments with CL and inspire more English teachers to make CL a regular component of their teaching repertoires.

Another important contribution of this research lies in the fact it provided a comparison between the CL approach and traditional teaching methods, which included evaluation of differences by means of a range of statistical analyses—t-tests, effect sizes, ANOVAs, ANCOVAs, and mean plots. It is considered that this combined use of a variety of statistical analysis has contributed to obtaining more reliable and accurate results. The comparison was also more comprehensive than previous studies because it included 28 specific areas within three dimensions: language competence, learning motivation and social skills.

In addition, based on an extensive literature review both on existing social skills assessments and classifications of CL skills, the researcher developed the SSS-CCEL with the aim of investigating the impact of CL on Chinese College English learners' social skills. Although some weaknesses exist in this seven-point Likert scale (which will be discussed in the following section on limitations of this research), there is no doubt that it has made up for the absence of applicable instruments in this field and has made the measuring of this area possible. In addition, the process of developing this measure, which is elaborated in the methodology chapter, may provide an insight into composing questionnaires for this purpose and somewhat lay the foundations for the future development of more reliable and valid measures of skills related to CL.

Limitations of this Research

In spite of above contributions, there are also a number of notable limitations in this research, which may affect the generalizability of its results. First, the experimental study included in this research was not a fully randomized control trial. The findings were based on comparison of two randomly assigned classes within a university in China, but individual students were not randomly selected for the intervention and comparison group, so there may have been some differences between the two groups. Also, this study did not include a true control group, which is not exposed to any intervention, either CL or traditional teaching methods, with the aim of being able to account for the influence of practice effects. This was not possible in the study because all

year-one students at the university are required to study English as a compulsory course. Further research, employing a full pre-test-post-test control group experimental design or involving a true control group, needs to be conducted to confirm the findings. Second, this experimental study was constrained to one semester of 18 weeks involving only two classes totalling 100 students within a university, whereas many millions of tertiary students study English for at least four semesters in China. Ideally, a longer experimental period on a larger scale involving more participants and universities would provide a more rigorous and accurate test of the impact of the intervention. Third, this study only focused on investigating learning outcomes from the two different teaching approaches. It did not include collection of data on the instructional processes which took place inside the classroom. There is increasing awareness that it is useful to incorporate both process and product data into the design of classroom research in order to have a comprehensive understanding of the effects of particular teaching approaches on learning (Nunan & Bailey, 2009). Fourth, the study did not include a follow-up to evaluate whether the positive impact of CL was maintained over time. Fifth, since the researcher, who believes in the potential of CL in EFL teaching, was the instructor of English for both groups, CL may not be as effective for other instructors with different pedagogical philosophies and teaching styles. Last, very few empirical studies have been conducted to compare effects between CL and traditional teaching on EFL learners' social skills and learning motivation. This lack of relevant literature and supporting evidence from previous studies makes it hard to compare the findings of this research with others' studies in the same field, and consequently limits validation of the findings of this research.

In addition, three factors—the use of self-report instruments, weaknesses in the SSS-CCEL, and the exclusive use of quantitative methodology—may affect the validity of this research. First, two instruments used in this research to respectively measure students' learning motivation and social skills—the LLOS and the SSS-CCEL—are in the form of self-report Likert scales. The information collected by means of self-report instruments usually gives a picture of respondents' belief about themselves rather than providing an accurate reflection of how they actually act. According to Gay et al. (2009, p. 153), "Self-report instruments such as attitude, interests, values, and personality scales have notable limits" due to "the existence of a response set, the tendency of

an individual to respond in a particular way to a variety of instruments”. People tend to select responses that are assumed to be more socially acceptable even if they are not actually their own characteristics and fit their own situations. In this research, although participants were informed that they were responding anonymously, there was still evidence that some responses reflected more of what respondents thought they were than what they actually were. For instance, when responding to item 15 “I feel stressed and uncomfortable when working with others in English classes”, both groups were found to have reported a high level of comfort and ease with group work at pre-test, especially the comparison group which registered a mean value approaching six on a seven-point scale. This finding was unlikely to be an accurate reflection of their true characteristics considering their EFL classroom teaching is usually dominated by teacher talk in China. So the use of self-report instruments may run the risk of yielding inaccurate data, and consequently may affect the validity of the research.

Furthermore, creating a well-established and high-quality measure requires considerable time, effort and expertise, as well as a long procedure of testing and re-testing, so the newly-developed SSS-CCEL may involve some weaknesses in spite of careful piloting beforehand. During the application of the SSS-CCEL in the main study of this research, some weaknesses have been located and are summarized as follows:

- 1) The internal consistency reliability is not very satisfactory because the Cronbach alpha coefficient of some subscales (e.g. self-confidence, sense of cohesion, initiative in socialization, and equal participation and accountability) are not consistently above the ideal level of 0.7;
- 2) Some items (e.g. “Luck decides most things that happen to me” and “Few people are as sensitive and understanding as I am”) are related to beliefs and dispositions which are generally stable and can hardly be changed by a short-term educational intervention in the EFL classroom. The inadequate relevance between some items and the specific intervention may somewhat decrease the sensitivity of the measure and affect its content validity. Some other items (e.g. “I offer teammates support and assistance as much as I can, so that they do their best in English” and “I believe everybody of different English levels can make a

contribution to the completion of group tasks”) could be considered somewhat biased in favour of CL due to the use of the terms “teammates” and “group tasks”, which may affect the fairness of outcomes.

- 3) Some possibly important variables were not properly reflected by the items on the measure. For instance, listening attentively is considered an essential collaborative skill throughout the whole process of peer interaction (Brown & Thomson, 2000; Jacobs et al., 2002; Jolliffe, 2007; Slavin, 2000) and very important for language learning (Ellis, 2009a), but no items were designed to represent this variable. This may, in some way, reduce the representativeness of the item samples, and create flaws in sampling validity—which refers to how well the test samples represent the total content areas being tested and is an important component of content validity (Gay et al., 2009).
- 4) The SSS-CCEL is intended to measure not only such specific skills as checking for understanding and initiative in socialization but also some complex attributes such as self-confidence, acceptance and empathy. For these more complex variables, a sample of five items may not be adequate to represent their potential aspects, so a wider range of items should be included for obtaining more valid and accurate data.

Lastly, this research is purely based on quantitative data without involvement of any interviews with either individual students or focus groups about their own perspectives on the use of CL and traditional instruction. In comparison with interview data, numerical data has the advantage of providing high-level generalizability as well as a broad understanding of the research setting, but has limitations regarding an in-depth exploration into a particular phenomenon of interest (Cohen et al., 2007; Gay et al., 2009). In this research, interpretations of the findings mainly relied on existing theories as well as the researcher’s perspectives, speculations and observations. These explanations sometimes appear to be insufficient since they lack the richness of students’ voices. So it would be ideal if follow-up interviews with some students had been conducted for the purpose of gaining a better understanding of this aspect. Considering the current trend of emphasising student-centred approaches and students’ needs, more emphasis should be placed on efforts to probe students’ feelings and perspectives on teaching innovations as well as on

their learning settings. Therefore, the lack of interview data with students forms another limitation of this research.

Recommendations for Future Research

A number of recommendations are provided with the aim of facilitating the design of rigorous research addressing the limitations of this study, and proposing important areas to investigate in future research. First, one prerequisite for conducting a rigorous comparative study between the CL approach and traditional instruction is the involvement of an adequate intervention in the actual process of teaching. Some studies (e.g. Sach et al., 2003) were considered somewhat inadequate in producing valid and reliable results because teachers were unable to apply CL activities into their classroom teaching on a frequent basis. Therefore, it is recommended that a careful pilot study should be implemented before the start of a formal experimental study. The pilot study provides teachers with time and opportunities to try out CL possibilities and figure out suitable techniques for their particular settings. Meanwhile, students have time and opportunities to develop the necessary skills for working together in teams. In other words, the pilot study may be greatly conducive to generating valid interventions in the subsequent main study and thus forms an integral part of a rigorous empirical study.

Second, in order to increase the generalizability and quality of studies, future researchers should pay attention to the following issues when conducting a comparative study between CL and traditional instruction in the context of College English teaching. These issues include: employing a full pre-test-post-test control group experimental design with the selection of participants on a random basis; extending the intervention period between pre-test and post-test to at least two semesters because a longer intervention tends to generate more reliable results; involving more universities and participants in the research programme which is conducive to promoting reliability of findings; incorporating both process data and data on learning outcomes; and including a follow-up aimed at evaluating the extent to which the impact of different teaching

approaches is maintained over time.

Third, in accordance with the inadequacies of the SSS-CCEL used in the current research, efforts should be made to work out a more suitable instrument to measure College English learners' cooperative skills, either through making modifications to the existing measure or through developing a new one. In the process of modifying or developing the measure, the main focus should be on a high level of internal reliability and content validity. In addition, special attention should be paid to increasing the appropriateness of the measure for the specific educational setting. For this purpose, emphasis should be placed on several issues: including items that are more relevant to the intervention as well as students' learning experiences, improving the inclusiveness of potential aspects of skills relevant to CL, including enough items to represent the broad aspects of a complex variable. Moreover, similar efforts should be made to improve the measure of motivation, for instance, through including items of more sensitivity and validity to detect the differences that a CL intervention can potentially bring to College English learners in China. It is recommended that cooperation with educational and social psychologists may help develop more reliable and valid measures particularly suitable for Chinese tertiary EFL learners.

The last recommendation is the use of a mixed methods research design including both quantitative data (e.g. from tests or surveys) and qualitative data (e.g. from interviews with students or teachers) in a single study. Mixed methods research is aimed at building on the "synergy and strength that exists between quantitative and qualitative research methods to understand a phenomenon more fully than is possible using either quantitative or qualitative methods alone" (Gay et al., 2009, p. 462). Obviously, this approach to research is greatly conducive to understanding research settings both broadly (i.e., from quantitative data) and deeply (i.e. from qualitative data). In addition, the in-depth and detailed qualitative data obtained from follow-up interviews with participants may partly remedy the possible effects of the response set (i.e. the tendency of an individual to select more socially-accepted responses) existing in self-report instruments, and thus may address the weakness in the use of survey questionnaires for quantitative data. Clearly, the use of mixed methods within in a single study can provide a more accurate and full-scale picture of participants' perspectives.

REFERENCES

- Abram, P., Scarloss, B., Holthuis, N., Cohen, E., Lotan, R., & Schultz, S. E. (2002). The use of evaluation criteria to improve academic discussion in cooperative groups. *Asia Pacific Journal of Education*, 22(1), 16-27.
- Aronson, E., & Patnoe, S. (1997). *The Jigsaw classroom: Building cooperation in the classroom* (2nd ed.). New York, NY: Longman.
- Baer, J. (2003) Grouping and achievement in cooperative learning. *College Teaching*, 51(4), 169-174.
- Bailey, K. M. (2003). Speaking. In D. Nunan (Ed.), *Practical English language teaching* (pp. 47-66). New York, NY: McGraw-Hill.
- Baloche, L. A. (1998). *The cooperative classroom: Empowering learning*. Upper Saddle River, NJ: Prentice Hall.
- Bejarano, Y. (1987). A cooperative small-group methodology in the language classroom. *TESOL Quarterly*, 21(3), 483-501.
- Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1997). *Women's ways of knowing: the development of self, voice, and mind* (Rev. ed). New York, NY: Basic Books.
- Bennett, N. (1996). Class size in primary schools: perceptions of headmasters, chairs of governors, teachers and parents. *British Educational Research Journal*, 22(1), 23-55.
- Bjorning-Gyde, M., Doogan, F., & East, M. (2008). Towards a fusion model for the teaching and learning of English in a Chinese context. In L. Dunn & M. Wallace (Eds.), *Teaching in transnational higher education: enhancing learning for offshore international students* (pp. 77-87). New York, NY: Routledge.
- Boud, D., Cohen, R., & Sampson, J. (2001). Peer learning and assessment. In D. Boud, R. Cohen & J. Sampson. (Eds.), *Peer learning in higher education: Learning from & with each other* (pp. 67-84). London: Kogan Page.
- Bracey, G. W. (2006). *Reading educational research: How to avoid getting statistically snookered*.

Portsmouth, NH: Heinemann.

- Brick, J. (1991). *China: A handbook in intercultural communication*. Sydney: National Centre for English Language Teaching and Research. .
- Brown, D., & Thomson, C. (2000). *Cooperative learning in New Zealand schools*. Palmerston North: Dunmore Press.
- Brown, G., & Atkins, M. (1991). *Effective teaching in higher education*. London: Routledge.
- Bruffee, K. A. (Ed.). (1993). *Collaborative learning: Higher education, interdependence and the authority of knowledge*. Baltimore, MD: Johns Hopkins University Press.
- Bruffee, K. A. (1995). Sharing our toys: cooperative learning versus collaborative learning. *Change*, 27(1), 12-18.
- Burden, P. (2000). The use of 'only English' in a learner-centred university classroom in Japan. *RELC Journal*, 31(1), 139-149.
- Campbell, D. T., & Stanley, J. C. (1971). *Experimental and quasi-experimental designs for research*. Chicago, IL: American Educational Research Association.
- Cannon, R., & Newble, D. (2000). *A handbook for teachers in universities and colleges: A guide to improving teaching methods* (4th ed.). London: Kogan Page.
- Chamot, A. U. & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Reading, MA: Addison-Wesley.
- Chen, H. (2007). Barriers in implementing the interactive teaching approach in language classes for non-English majors. *US-China Foreign Language*, 5(4), 29-35.
- Chen, H. (2008). On American cooperative language learning approach. *Educational Practice and Research*, 10, 45-47.
- Chen, M. L. (2005). *The effects of the cooperative learning instructional approach on Taiwanese EFL students' motivation, English listening, reading, and speaking competencies* (Unpublished doctoral dissertation). La Sierra University, Riverside, California, United States of America.
- Chen, R., & Hird, B. (2006). How do Chinese students collaborate in EFL group work? *Prospect*, 21(2), 70-81.

- Chin, L. F., Teh, P. L., & Fong, H. K. (1988). The effectiveness of cooperative and individualistic approaches in teaching mathematics and English. *Singapore Journal of Education*, 9(2), 1-17.
- Clarke, J. (1994). Pieces of the puzzle: The jigsaw method. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 34-50). Westport, CT: Greenwood Press.
- Clement, R., Dörnyei, Z., & Noels, K. (1994). Motivation, self-confidence and group cohesion in the foreign language classroom. *Language Learning*, 44(3), 417-448.
- Cohen, E. G. (1994). *Designing group work: strategies for the heterogeneous classroom* (2nd ed.). New York, NY: Teachers College Press.
- Cohen, E. G., Brody, C. M., & Sapon-Shevin, M. (Eds.) (2004). *Teaching cooperative learning: The challenge for teacher education*. Albany, NY: State University of New York Press.
- Cohen, E. G., Lotan, R. A., Whitcomb, J. A., Balderrama, M. V., Cossey, R. & Swanson, P. E. (1994). Complex Instruction: Higher-order thinking in heterogeneous classrooms. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 82-96), Westport, CT: Greenwood Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). London: Routledge.
- Comanaru, R., & Noels, K. A. (2009). Self-determination, motivation, and the learning of Chinese as a heritage language. *The Canadian Modern Language Review*, 66(1), 131-158.
- DaSilva Iddings, A. C. (2006). Conclusion. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 177-180). Cambridge: Cambridge University Press.
- DeCarvalho, R. (1991). The humanistic paradigm in education. *The Humanistic Psychologist*, 19(1), 88-104.
- Demaray, M. K., & Ruffalo, S. L. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, 24(4), 648-671.

- Deng, X. (2007). Promotion of interaction in cooperative learning tasks. *Sino-US English Teaching*, 4(7), 8-13.
- Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations*, 2(2), 129-152.
- DeVellis, R. F. (2003). *Scale development: theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- Dewey, J. (1924). *The school and society*. Chicago, CL: University of Chicago Press.
- Dishon, D., & O'Leary, P. W. (1998). *Guidebook for cooperative learning: Techniques for creating more effective schools* (3rd ed.). Holmes Beach, FL: Learning Publications.
- Dörnyei, Z. (1997). Psychological Processes in cooperative language learning: Group dynamics and motivation. *The Modern Language Journal*, 81(4), 482-493.
- Dörnyei, Z. (2006). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Drew, C. J., Hardman, M. L., & Hosp, J. L. (2008). *Designing and conducting research in education*. Thousand Oaks, CA: Sage.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam: John Benjamin.
- Ellis, R. (2009a). A reader responds to Guilloteaux and Dörnyei's "Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation". *TESOL Quarterly*, 43(1), 105-109.
- Ellis, R. (2009b). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(2), 221-246.
- Ellis, R. (2009c). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474-509.
- Falsgraf, C. (2009). The ecology of assessment. *Language Teaching*, 42(4), 491-503.
- Flowerdew, L. (1998). A cultural perspective on group work. *ELT Journal*, 52(4), 323-329.
- Freeman, D. E., & Freeman, Y. S. (1994). *Between worlds: Access to second language acquisition*. Portsmouth, NH: Heinemann.
- Fullan, M. G., Bennett, B., & Rolheiser-Bennett, C. (1990). *Linking classroom and school*

- improvement. *Educational Leadership*, 47(8), 13-19.
- Gage, N., & Berliner, D. (1998). *Educational psychology* (6th ed.). Boston, MA: Houghton Mifflin Harcourt.
- Gagne, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26, 331-362.
- Galassi, J. P., DeLo, J. S., Galassi, M. D., & Bastien, S. (1974). The college self-expression scale: A measure of assertiveness. *Behavior Therapy*, 5(2), 165-171.
- Garson, G. D. (2008). GLM Repeated measures. Retrieved November 13, 2009, from <http://faculty.chass.ncsu.edu/garson/PA765/glmrepeated.htm>
- Garson, G. D. (2009). Univariate GLM, ANOVA, and ANCOVA. Retrieved November 13, 2009, from <http://faculty.chass.ncsu.edu/garson/PA765/anova.htm#assume>
- Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications* (9th ed.). London: Pearson Education.
- Gay, M. L., Hollandsworth, J. G., & Galassi, J. P. (1975). An assertiveness inventory for adults. *Journal of Counselling Psychology*, 22(4), 340-344.
- Ghaith, G. (2003). Effects of the Learning Together Model of cooperative learning on English as a foreign language reading achievement, academic self-esteem, and feelings of school alienation. *Bilingual Research Journal*, 27(3), 451-475.
- Gibbs, G., & Habeshaw, T. (1989). *Preparing to teach: An introduction to effective teaching in higher education*. Bristol: Technical and Educational Services.
- Gillies, R. M. (2007). *Cooperative learning: Integrating theory and practice*. Thousand Oaks, CA: Sage.
- Gillies, R. M., & Ashman, A. F. (Eds.). (2003a). *The social and intellectual outcomes of learning in groups*. London: RoutledgeFalmer.
- Gillies, R. M., & Ashman, A. F. (2003b). An historical review of the use of groups to promote socialization and learning. In R. M. Gillies & A. F. Ashman (Eds.), *The social and intellectual outcomes of learning in groups* (pp. 1-18). London: RoutledgeFalmer.
- Gömleksiz, M. N. (2007). Effectiveness of cooperative learning (Jigsaw II) method in teaching

- English as a foreign language to engineering students (Case of Firat University, Turkey). *European Journal of Engineering Education*. 32(5), 613-624.
- Guo, C. (2001, August 13). Looking at China from two hot issues. *People's Daily Overseas Edition*, p. 1.
- Guo X. (2004). Cooperative learning in large classes. *CELEA Journal*, 27(4), 43-46.
- Gwyn-Paquette, C., & Tochon, F. V. (2002). The role of reflective conversations and feedback in helping pre-service teachers learn to use cooperative activities in their second language classrooms. *The Modern Language Journal*, 86(2), 204-226.
- Harmer, J. (1998). *How to teach English: An introduction to the practice of English language teaching*. Harlow: Pearson Education.
- Harmer, J. (2007). *The practice of English language teaching* (2nd ed.). Harlow: Pearson Education.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- He, M. (2003). Strategies for promoting cooperative learning in EFL classes in China. *Teaching English in China*, 26(4), 116-121.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- High, J. (1993). *Second language learning through cooperative learning*. San Juan Capistrano, CA: Kagan Cooperative Learning.
- Hill, S., & Eckert, P. (1995). *Leading communities of learners*. Salisbury East: Management and Research Centre.
- Hofstede, G. (1986). Cultural difference in teaching and learning. *International Journal of Intercultural Relations*, 10(3), 301-320.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.
- Holliday, A. (1994). *Appropriate methodology and social context*. Cambridge: Cambridge University Press.

- Holt, D. (Ed.) (1993). Cooperative learning: A response to linguistic and cultural diversity. McHenry, IL: Centre for Applied Linguistics.
- Hornby, G. (2009). The effectiveness of cooperative learning with trainee teachers. *Journal of Education for Teaching*, 35(2), 161-168.
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Cultural and Curriculum*, 15(2), 93-105.
- Hu, Z. (2005). China's English education reform: Trends and issues. *Foreign language in China*, 6(2), 4-10.
- Inderbitzen, H. M., & Foster, S. L. (1992). The teenage inventory of social skills: Development, reliability, and validity. *Psychological Assessment*, 4(4), 451-459.
- Jacobs, G. M. (2006). Issues in implementing cooperative learning. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 20-46). Cambridge: Cambridge University Press.
- Jacobs, G. M., & Goh, C. C. (2007). *Cooperative learning in the language classroom*. Singapore: SEAMEO Regional Language Centre.
- Jacobs, G. M., Lee, G. S., & Ball, J. (1996). *Learning cooperative learning via cooperative learning: A sourcebook of lesson plans for teacher education on cooperative learning*. Singapore: SEAMEO Regional Language Centre.
- Jacobs, G. M., & McCafferty, S. G. (2006). Connections between cooperative learning and second language learning and teaching. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 18-29). Cambridge: Cambridge University Press.
- Jacobs, G. M., McCafferty, S. G., & DaSilva Iddings, A. C. (2006). Roots of cooperative learning in general education. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 9-17). Cambridge: Cambridge University Press.
- Jacobs, G. M., Power, M. A., & Inn, L. W. (2002). *The teacher's sourcebook for cooperative learning: Practical techniques, basic principles, and frequently asked questions*.

Thousand Oaks, CA: Corwin Press.

- Jacobs, G. M., & Ratmanida. (1996). The appropriacy of group activities: Views from some Southeast Asian second language education. *RELC Journal*, 27(1), 103-120.
- Jin, L., & Cortazzi, M. (2004). English language teaching in China: A bridge to the future. In W. K. Ho & R. Y. Wong (Eds.), *English language teaching in East Asia today: Changing policies and practices* (pp. 119-134). Singapore: Eastern Universities Press.
- Johnson, D. W., & Johnson, R. T. (1974). Instructional goal structure: Cooperative, competitive, or individualistic. *Review of Educational Research*, 44(2), 213-240.
- Johnson, D. W., & Johnson, R. T. (1992). *Positive interdependence: The heart of cooperative learning*. Edina, MN: Interaction Book.
- Johnson, D. W., & Johnson, R. T. (1994a). Learning together. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 51-65). Westport, CT: Greenwood Press.
- Johnson, D. W., & Johnson, R. T. (1994b). *Learning together and alone* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Johnson, D. W., & Johnson, R. T. (1994c). Structuring academic controversy. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 66-81). Westport, CT: Greenwood Press.
- Johnson, D. W., & Johnson, R. T. (2003). Student motivation in co-operative groups: Social interdependence theory. In R. M. Gillies & A. F. Ashman (Eds.), *Cooperative learning: The social and intellectual outcomes of learning in groups* (pp. 137-176). London: RoutledgeFalmer.
- Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1994). *The new circles of learning: Cooperation in the classroom and school*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1998). *Cooperation in the classroom* (7th ed.). Edina, MN: Interaction Book.
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, 19(1), 15-29.
- Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D., & Skon, L. (1981). Effects of

- cooperative, competitive and individualistic goal structures on achievement: a meta-analysis. *Psychological Bulletin*, 89(1), 47-62.
- Jolliffe, W. (2007). *Cooperative learning in the classroom: Putting it into practice*. London: Paul Chapman.
- Joritz-Nakagawa, J. (2006). Integrating global education and cooperative learning in a university foreign language reading class. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 134-152). Cambridge: Cambridge University Press.
- Kagan, S. (1994). *Cooperative learning*. San Clemente, CA: Kagan Cooperative Learning.
- Kagan, S., & Kagan, M. (1994). The Structural Approach: Six keys to cooperative learning. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 115-133), Westport, CT: Greenwood Press.
- Kagan, S., & Kagan, M. (2009). *Kagan cooperative learning*. San Clemente, CA: Kagan Publishing.
- Kell, J., & Newton, C. (1997). Roles of pathways in self-access centres. *ELT Journal*, 51(1), 48-53.
- Kerlinger, F. N. (1970). *Foundations of behavioral research*. New York, NY: Holt, Rinehart & Winston.
- Kramsch, C., & Sullivan, P. (1996). Appropriate pedagogy. *ELT Journal*, 50 (3), 199-212.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. London: Routledge.
- Le, H. (2006). Learning vocabulary in group work in Vietnam. *RELC Journal*, 37(1), 105-121.
- Lewin, K. (1935). *A dynamic theory of personality*. New York, NY: McGraw-Hill.
- Lewin, K. (1948). *Resolving social conflicts*. New York, NY: Harper.
- Li, J. (2007). An experimental study on the feasibility of applying cooperative learning to College English large classes. *Journal of Xi'an International Studies University*, 15(3), 91-94.

- Li, K. (2008). Teaching College English to a large class: Problems and remedies. *Sino-US English Teaching*, 5(2), 13-18.
- Li, L. (1996, September 5). Vice-premier Lanqing Li's speech at the national conference on foreign language education. *Wenhuibao*, p. 1.
- Li, M., & Campbell, J. (2008). Asian students' perceptions of group work and group assignments in a New Zealand tertiary institution. *Intercultural Education*, 19(3), 203-216.
- Li, R. (2007). *Reconsidering the application of cooperative learning in College English classroom teaching* (Unpublished master's thesis). Shanghai International Studies University, Shanghai, China.
- Li, R., Dong, T., & Ma, J. (2006). On the model of computer and network based cooperative autonomous learning. *Journal of Hebei University*, 31(4), 126-129.
- Li, X. (2007). On network-based cooperative learning in College English teaching. *Journal of Henan Normal University*, 34(5), 233-235.
- Liang, X. (2004). Cooperative learning as a sociocultural practice. *The Canadian Modern Language Review*, 60(5), 637-668.
- Liao, X. (2004). The need for communicative language teaching in China. *ELT Journal*, 58(3), 270-273.
- Littlewood, W. (2000). Do Asian students really want to listen and obey? *ELT Journal*, 54(1), 31-36.
- Liu, J. (2002). Negotiating silence in American classrooms: Three Chinese cases. *Language and Intercultural Communication*, 2(1), 37-54.
- Liu, R., & Dai, M. (2004). On the reform of College English teaching in China. *CELEA Journal*, 27(4), 3-8.
- Liu, S. (2007). On cooperative language learning in the College English class. *Journal of Shenyang Normal University*, 31(4), 162-164.
- Long, M. H. (1983). Native-speaker/non-native speaker conversations and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126-141.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W.

- C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207-228.
- Lorr, M., Youniss, R. P., & Stefic E. C. (1991). An inventory of social skills. *Journal of Personality Assessment*, 57(3), 506-520.
- Luo, J. (2007). Cooperative learning in the English classroom: Application and reflection. *Journal of Basic English Education*, 9(1), 62-63.
- Magee, V. Y. G., & Jacobs, G. M. (2001). Comparing second language student participation under three teaching modes. *Journal of the Chinese Language Teachers Association*, 36(1), 61-80.
- Matthews, R. S., Cooper, J. L., Davidson, N., & Hawkes, P. (1995). Building bridges between cooperative and collaborative learning. *Change*, 27(1), 35-40.
- McCafferty, S. G., Jacobs, G. M., & DaSilva Iddings, A. C. (Eds.). (2006). *Cooperative learning and second language teaching*. Cambridge: Cambridge University Press.
- McGroarty, M. (1993). Cooperative learning and second language acquisition. In D. D. Holt (Ed.), *Cooperative learning: A response to linguistic and cultural diversity* (pp. 19-46). McHenry, IL: Centre for Applied Linguistics.
- McGuire, S. P. (1992). *An application of cooperative learning to teaching English as a foreign language in Japan* (Unpublished master's thesis). University of Minnesota, Minnesota, United States.
- McWhaw, K., Schnackenberg, H., Sclater, J., & Abrami, P. (2003). From co-operation to collaboration: Helping students become collaborative learners. In R. M. Gillies & A. F. Ashman (Eds.), *The social and intellectual outcomes of learning in groups* (pp. 1-18). London: RoutledgeFalmer.
- Miller, L., & Aldred, D. (2000). Student teachers' perceptions about communicative language teaching. *RELC Journal*, 31(1), 1-22.
- Ministry of Education. (2001). *Some opinions on strengthening English teaching for*

- undergraduates and improving teaching quality in higher education*. Beijing: Author.
- Ministry of Education. (2004). *College English Curriculum Requirements for Trial Implementation*. Beijing: Foreign Language Teaching and Research Press.
- Ministry of Education. (2007). *College English Curriculum Requirements*. Shanghai: Foreign Language Education Press.
- Murdoch, K., & Wilson, J. (2008). *Helping your pupils to work cooperatively*. London: Routledge.
- Mutch, C. (2005). *Doing educational research: A practitioner's guide to getting started*. Wellington: New Zealand Council for Educational Research Press.
- Ning, H. (2010). Adapting cooperative learning in tertiary ELT. *ELT Journal*, doi: 10.1093/elt/ccq021.
- Ning, H., & Hornby, G. (2010). Effectiveness of cooperative learning in teaching English to Chinese tertiary learners. *Effective Education*, 2(2), 99-116.
- Noels, K. A., Pelletier, L. G., Clement, R., & Vallerand R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 50(1), 57-85.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, doi: 10.1007/s10459-010-9222-y.
- Nunan, D. (Ed.) (2003a). *Practical English language teaching*. New York: McGraw-Hill.
- Nunan, D. (2003b). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, 37(4), 589-613.
- Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, MA: Heinle.
- Oxford, R. L. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal*, 81(4), 443-456.
- Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS version 15* (3rd ed.). Maidenhead: Open University Press.
- Park, C. C. (2002). Cross-cultural differences in learning style of secondary English learner.

Bilingual Research Journal, 26(2), 443-459.

- Peng, C., & Qi, A. (2006). An experimental research on the cooperative learning mode in College English teaching. *Journal of Nanchang Institute of Aeronautical Technology*, 8(1), 91-95.
- Peng, J. E. (2007). Willingness to communicate in the Chinese EFL classroom: A cultural perspective. In J. Liu (Ed.), *English language teaching in China: New approaches, perspectives and standards* (pp. 250-269). London: Continuum International.
- Phipps, M., Phipps, C., Kask, S., & Higgins, S. (2001). University students' perceptions of cooperative learning: Implications for administrators and instructors. *The Journal of Experiential Education*, 24(1), 14-21.
- Phuong-Mai, N., Terlouw, C., & Pilot, A. (2006). Culturally appropriate pedagogy: the case of group learning in a Confucian Heritage Culture context. *Intercultural Education*, 17(1), 1-19.
- Piaget, J. (1959). *The language and thought of the child* (3rd ed.). London: Routledge & Kegan-Paul.
- Pica, T., & Doughty, C. (1985). The role of group work in classroom second language acquisition. *Studies in Second Language Acquisition*, 7(2), 233-248.
- Powell, B. (2009, November 23). 5 things the U.S. can learn from China. *Time*, 174(20), 24-31.
- Reid, J. M. (1993). *Teaching ESL writing*. Englewood Cliffs, NJ: Prentice Hall.
- Richards, J. & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, 51(3), 649-660.
- Riggio, R. E. (2003). The social skills inventory (SSI): Measuring nonverbal and social skills. In V. Manusov (Ed.), *The source book of nonverbal measures: Going beyond words* (pp. 25-33). Mahwah, NJ: Lawrence Erlbaum Associates.
- Riggio, R. E., & Carney, D. R. (2003). *Social skills inventory manual* (2nd ed.). Retrieved May 29, 2007, from http://www.mindgarden.info/files/MGI_NING_SSIN9269.pdf
- Riordan, D., Street, D., & Roof, B. (Eds.) (1997). *Group learning: Application in higher education*.

- Harrisonburg, VA: James Madison University, Institute for Research in Higher Education.
- Roberts, T. S. (Ed.). (2004). *Online collaborative learning: Theory and practice*. Hershey, PA: Information Science.
- Roberts, T. S. (Ed.). (2005). *Computer-supported collaborative learning in higher education*. Hershey, PA: Idea Group.
- Robson, C. (1993). *Real world research: A resource for social scientists and practitioner-researchers*. Cambridge, MA: Blackwell.
- Rogers, C. R., & Freiberg, H. J. (1994). *Freedom to Learn* (3rd ed.). Upper Saddle River, NJ: Merrill.
- Romney, J. C. (1997). Collaborative learning in a translation course. *Canadian Modern Language Review*, 54, 48-67.
- Ruan, Y., & Jacob, W. J. (2009). The transformation of College English in China. *Frontier of Education in China*, 4(3), 466-487.
- Sachs, G. T., Candlin, C. N., Rose, K. R. & Shum, S. (2003). Developing cooperative learning in the EFL/ESL secondary classroom. *RELJ Journal*, 34(3), 338-369.
- Sharan, S. (Ed.) (1994). *Handbook of cooperative learning methods*. Westport, CT: Greenwood Press.
- Sharan, S. (2002). Differentiating methods of cooperative learning in research and practice. *Asia Pacific Journal of Education*, 22(1), 106-116.
- Sharan, S. (2003). Large classes, small groups: A social systems approach. In R. M. Gillies & A. F. Ashman (Eds.), *The social and intellectual outcomes of learning in groups* (pp. 1-18). London: RoutledgeFalmer.
- Sharan, S. & Shachar, H. (1988). *Language and learning in the cooperative classroom*. New York: Springer-Verlag.
- Sharan, Y. & Sharan, S. (1994). Group Investigation in the cooperative classroom. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 97-114), Westport, CT: Greenwood Press.
- Shu, D. (2004). *FLT in China: Problems and suggested solutions*. Shanghai: Shanghai Foreign

Language Education Press.

- Siemon, A. (2010). Learning English in China: Is communicative language teaching a reality? *Practically Primary*, 15(2), 39-42.
- Slavin, R. E. (1987). Cooperative learning: where behavioural and humanistic approaches to classroom motivation meet. *The Elementary School Journal*, 88(1), 29-37.
- Slavin, R. E. (1990). IBM's writing to read: Is it right for reading? *Phi Delta Kappan*, 72(3), 214-216.
- Slavin, R. E. (1994). Student Teams-Achievement Division. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 3-19), Westport, CT: Greenwood Press.
- Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Needham Heights, MA: Allyn and Bacon.
- Slavin, R. E. (1997). Research on cooperative learning: Consensus and controversy. In D. Riordan, D. Street & B. Roof (Eds.), *Group learning: Applications in higher education* (pp. 3-7). Harrisonburg, VA: James Madison University, Institute for Research in Higher Education.
- Slavin, R. E. (2000). *Educational Psychology: Theory and practice* (6th ed.). Boston, MA: Allyn and Bacon.
- Stone, J. (1994). *Cooperative learning and language arts: A multi-structural approach*. San Juan Capistrano, CA: Kagan Cooperative Learning.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50(1), 158-164.
- Tang, S., Zhang, L. J., & Dong Y. (2009). Integrating cooperative learning into genre-based teaching of EFL writing. *CELEA Journal*, 32(1), 99-108.
- Teng, H., Niu, Q. & Wolff, M. (2004). China's ESL goals: Are they being met?' *English Today*, 20(3), 37-44.
- Tjosvold, D., & Fang, S. S. (2004). Cooperative conflict management as a basis for training

- students in China. *Theory into Practice*, 43(1), 80-86.
- Tjosvold, D., Nibler, R., & Wan, P. (2001). Motivation for conflict among Chinese university students: Effect of others' expertise and one's own confidence on engaging in conflict. *The Journal of Social Psychology*, 14(3), 353-363.
- Totten, S., Sills, T., Digby, A., & Russ, P. (1991). *Cooperative learning: A guide to research*. New York, NY: Garland Publishing.
- Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, 6(1), 71-90.
- Vella, J. (2002). *Learning to listen, learning to teach: The power of dialogue in educating adults* (Rev. ed.). San Francisco, CA: Jossey-Bass.
- Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Eds.). Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Wang, Y. (1999). College English in China. *English Today*, 15(1), 45-51.
- Wang, Z. (2004). *A general study on cooperative learning theory and its application to oral English teaching* (Unpublished master's dissertation). Capital Normal University, Beijing, China.
- Warschauer, M., & Kern, R. (Eds.). (2000). *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.
- Waugh, R. F., Bowering, M. H. & Chayarathee, S. (2005). Cooperative learning versus communicative Thai teaching of English as a second language for Prathom (Grade) 6 students taught in Thailand. In R. F. Waugh (Ed.), *Frontiers in educational psychology* (pp. 221-232). Hauppauge, NY: Nova Science.
- Wee, S. & Jacob G. M. (2006). Implementing cooperative learning with secondary school students. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 113-133). Cambridge: Cambridge University Press.
- Wette, R. & Barkhuizen, G. (2009). Teaching the book and educating the person: challenges for

- university English language teachers in China. *Asia Pacific Journal of Education*, 29(2), 195-212.
- Wilhelm, K. H. (2006). Teaching, practicing, and celebrating a cooperative learning model. In S. G. McCafferty, G. M. Jacobs & A. C. DaSilva Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 153-176). Cambridge: Cambridge University Press.
- Wittrock, M. (1990). Generative process of comprehension. *Educational Psychologist*, 24, 345-376.
- Wu, Y. (2001). English language teaching in China: Trends and challenges. *TESOL Quarterly*, 35(1), 191-194.
- Xia, J. (2002). A research report on College English teachers' educational perceptions, professional knowledge and competence, research and training programmes. *Foreign Language World*, (5), 35-41.
- Yang, H. & Weir, C. (1998). *Validation study of the National College English Test*. Shanghai: Shanghai Foreign Language Education Press.
- Yang, Y., & Nicholas, H. (2008). A review of 'interest in the learning of English' in the Chinese context. *The Asian EFL Journal*, 10(3), 146-156.
- Zhang, F., & Zhao, T. (2004). The application of cooperative learning theories into College English reading classes. *Foreign Language World*, (6), 46-51.
- Zhang, L. (2006). Preferred teaching styles and modes of thinking among university students in mainland China. *Thinking Skills and Creativity*, 1(2), 95-107.
- Zhang, Y. (2007). *Cooperative learning and College English listening teaching* (Unpublished master's dissertation). Shanghai International Studies University, Shanghai, China.

Appendix 1: A Summary of teaching content of College English course for reading and writing covered in the intervention

Unit	Section	Title	Learning objectives
1	A	Time-conscious American	1. Understanding the texts; 2. Learning new vocabulary (e.g. restless, marvellous, result in, worth doing, account for); 3. Reading skill—reading for main ideas in a passage; 4. Writing skill—providing details (examples, a list of items) to support a general statement.
	B	Culture shock	
2	A	Environmental protection throughout the world	1. Understanding the texts; 2. Learning new vocabulary (e.g. global, extensive, cod, contamination, strengthen, as a result of); 3. Reading skill—working out the meaning of unfamiliar words from context clues; 4. Writing skill—a basic writing structure on problem-and-solution
	B	Green spaces in cities	
3	A	Marriage across the nations	1. Understanding the texts; 2. Learning new vocabulary (e.g. intermarriage, tolerance, compromise, prejudice, cabbage, suspect, indicate, ups and downs); 3. Reading skill—recognizing differences between facts and opinions, and reading between lines; 4. Writing skill—presenting ideas (opinions, or arguments) and then giving reasons to accept, correct or reject the idea (opinions, or arguments).
	B	Rich meeting his future mother-in-law	
4	A	Study abroad	1. Understanding the texts; 2. Learning new vocabulary (e.g. certificate, remarkable, fluent, overseas, adapt, depend on, in turn, after all); 3. Reading skill—reading for the main idea in a sentence; 4. Writing skill—providing reasons to support a general statement.
	B	Experiences in Exile	

(continued)

Appendix 1: Continued

Unit	Section	Title	Learning objectives
5	A	Weeping for my smoking daughter	<ol style="list-style-type: none"> 1. Understanding the texts; 2. Learning new vocabulary (e.g. weep, spoil, harden, unnoticeable, slim, self-poisoning, insensitive, filter, die of); 3. Reading skill—understanding figurative language; 4. Writing skill—a basic writing structure on cause-and-effect.
	B	Stop spoiling your children	
6	A	As his name is, so is he	<ol style="list-style-type: none"> 1. Understanding the texts; 2. Learning new vocabulary (e.g. prominent, impressive, quality, award, turn off, turn down, stick to); 3. Reading skill—reading for the main idea in a paragraph; 4. Writing skill—using comparison to support an opinion or statement.
	B	Judge by appearance	
7	A	Lighten your load and save your life	<ol style="list-style-type: none"> 1. Understanding the texts; 2. Learning new vocabulary (e.g. lighten, overwhelm, stimulate, priority, workaholic, out of control, burn the midnight oil); 3. Reading skill—using context clues to work out word meanings; 4. Writing skill—a basic writing structure on cause-and-effect (cause)-and-effect (the previous effect becomes the cause of the next effect).
	B	Are you a workaholic?	
8	A	There's a lot more to life than a job	<ol style="list-style-type: none"> 1. Understanding the texts; 2. Learning new vocabulary (e.g. consumerist, incredible, constructive, distinguish, by comparison, in the long run, no wonder); 3. Reading skill—distinguishing between facts and opinions; 4. Writing skill—presenting an argument.
	B	What youngsters expect in life	
9	A	Stop brain drain	<ol style="list-style-type: none"> 1. Understanding the texts; 2. Learning new vocabulary (e.g. intermediate, commercial, parallel, mushroom, integrate, drain of, get away with, hold water); 3. Reading skill—reading for detailed information; 4. Writing skill—making contrasts or comparisons between two things, and practising the writing structure on cause-and-effect.
	B	Borderline ridiculousness	

Appendix 2: Summary of teaching content of College English course for listening and speaking covered in the intervention

Unit	Title	Learning objectives
1	Roll over, Beethoven!	<ol style="list-style-type: none"> 1. Making inference; 2. Expressing likes and dislikes; 3. Making suggestions.
2	What's on at the theatre?	<ol style="list-style-type: none"> 1. Listening for details; 2. Making comments; 3. Giving and accepting invitations.
3	Every Jack has his Jill!	<ol style="list-style-type: none"> 1. Identifying people or the relationship between the speakers; 2. Comforting people; 3. Expressing worries.
4	Beware of ads!	<ol style="list-style-type: none"> 1. Identifying the meaning of advertisements; 2. Clarifying and asking for clarification; 3. Speaking elliptically.
5	Does your best friend have four legs?	<ol style="list-style-type: none"> 1. Identifying special quantities; 2. Making recommendations; 3. Giving praise.
6	What's in fashion?	<ol style="list-style-type: none"> 1. Identifying the relationship between the speakers; 2. Asking for opinions; 3. Giving opinions.
7	Does money talk?	<ol style="list-style-type: none"> 1. Identifying larger numbers; 2. Expressing real and unreal conditions.
8	Crime does pay!	<ol style="list-style-type: none"> 1. Identifying signal words; 2. Bringing up sensitive topics; 3. Asking for details.

Appendix 3: A sample assignment of a team-task involving elements of Jigsaw, Complex Instruction and Group Investigation

Team name _____

Deadline: June 30th (a week away) for the first four questions
 July 9th (over two weeks away) for the fifth question

Task division for individual members:

Apple:

Bean:

Cat:

Dog:

Self-grading on Question 1 to 4 _____; Question 5 _____

Grade of the presenting teams:

Question	Team Name	Grade	Team Name	Grade
Ques 1-3				
Ques 4				
Ques 1-4 (on average)				
Ques 5				

Task requirements:

1. Rewrite “stealing brains from the third world” with the phrase of “drain ... of...” (line 4);
2. Analyse and paraphrase the long sentence from line 38 to 41;
3. Why are those foreign nationals called “temporary skilled personnel” (line 11);
4. Considering the situation mentioned in paragraph five about India, have we seen some parallels between China and India? No matter whether your answer is Yes or No, please give examples within three minutes to support your answers;
5. Prepare a ten-minute speech on “the brain drain in developing countries: reasons and results”. This task involves your comprehension of the text, your perspectives on the issue and after-class cooperative research. The team work will be assessed in the language laboratory, so you are welcome to use some visual aids such as pictures, PPT, or videos (the length of videos

is no more than two minutes), which can well improve the level of interestingness and comprehensibility of your speech.

Grading Criteria:

1. The first four questions are graded together on June 30th (next Monday) through the following procedures:
 - a. Two teams are randomly selected.
 - b. Then a representative is chosen from each team to put on the blackboard their answers to the three questions; meanwhile, one representative from each team answers the fourth question orally.
 - c. The first three questions are graded according to the accuracy of answers.
 - d. The fourth question is graded based on our commonly-used criteria for oral presentations, concerning fluency, loudness, clarity, comprehensibility, coherence, and length of their speech, as well as the level of interestingness.
 - e. Two grades (respectively for the first three questions and the fourth one) for each team average out to a general grade for each team.
2. The fifth question is graded separately on July 9th (Wednesday in two weeks). Apart from the points mentioned above in (d), grading criteria also relate to abilities in using computer technology to produce relevant visual aids for the oral presentation.
3. Eye contact is needed in the oral presentation.
4. Your team is attentive and quiet when others are speaking.

Scaffolding phrases and expressions:

1. It is obvious/clear/apparent/undisputed that there are some parallels/there is parallelism between China and India regarding the problem of the brain drain.
2. These days we can see growing differences between China and India relating to the situation mentioned in paragraph five.
3. The similarities/differences between the two nations can be illustrated through a number of examples. First of all/In addition/Moreover...
4. Signal words for reasons: due to, owing to, because of, because, since, for, result from...
5. Signal words for results: in consequence, as a consequence, consequently, therefore, so, thus, as a result, result in...

Appendix 4: An example of a worksheet for teaching reading

Team name _____

Time limit: 20 minutes

Individual roles: Recorder _____

Time monitor _____

Understanding checker _____

Taskmaster _____

Self-grading _____

Grade of the presenting teams:

Team name	Grade	Team name	Grade

Task requirements:

Read through the text and have a team discussion about which country the narrator originally comes from. Please give reasons.

Apple: (opinions from Apple)

Bean:

Cat:

Dog:

Team conclusion and supporting details:

Grading criteria:

1. You speak loud enough to be heard by all;
2. Your ideas are clearly expressed, well organized and easy to follow;
3. The length of your speech is no less than two minutes;
4. Eye contact is needed (You may use notes but shouldn't read);
5. Your team is attentive and quiet when others are speaking.

Scaffolding phrases and expressions:

1. It is obvious/clear/apparent/evident/ that the narrator originally comes from ... This is because, first/second/third...
2. There are also some evidence/context clues that indicate/suggest/show that the narrator is originally from For instance, ...
3. Sorry, I didn't get what you mean. Would you say it again?
4. Thank you, Apple. That's a great point. What do you think, Bean?
5. Let's focus on the topic. We're a bit off track.
6. Come on, guys. Only two minutes left.

Appendix 5: A team-grade sheet used by teams for peer- and self-grading

Assessing team's name _____

Grade range: A⁺, A, A⁻, B⁺, B, B⁻, C⁺ and C

time team grade name	2 nd week		3 rd week		4 th week		5 th week		6 th week		7 th week	
	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed
Stars												
Noise Maker												
Green Apple												
Backstreet												
X-pub												
Cuba												
Law-V												
Friends												
hotpot												
Fun												
Sunrise												
Hustle												
Beam												

Appendix 6: A teamwork recording sheet for the teacher's use

time grade name	2 nd week		3 rd week		4 th week		5 th week		6 th week		7 th week	
	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed	Mon	Wed
<u>BP</u>	2			2		1		2				1
Stars	C⁺		B⁻	B		A⁻			A			
Apple	C											
Bean				B								B ⁺
Cat			B ⁻									
Dog						B ⁺			A			
<u>BP</u>				1	2		1			2		
Noise maker		B		A⁻			A		B			
Apple									B ⁻			
Bean				B ⁺							A ⁻	
Cat		B										
Dog							A					
<u>BP</u>												
Green Apple												
Apple												
Bean												
Cat												
Dog												
<u>BP</u>												
Backstreet												
Apple												
Bean												
Cat												
Dog												
<u>BP</u>												
X-pub												
Apple												
Bean												
Cat												
Dog												

Note: 1. BP = Bonus point

2. The grades recorded after the team name may have been adjusted according to the extent of improvement.

Appendix 7: Contrasts of a typical 50-minute session between the two different methods for teaching listening and speaking

Unit title: Does your best friend have four legs?	
I. Warm-up questions: What do you think we are going to talk about from the title of this unit? What do “four-legged friends” refer to? Are there any exceptions? (five minutes)	
The intervention group	The comparison group
<ol style="list-style-type: none"> 1. Students worked in their teams for two minutes; 2. A student was randomly selected to represent his/her team to present answers; 3. The presenter’s performance was assessed by his/her home team and audience teams, and each team put their agreed-grade on the worksheet or team-grade sheet which had been collected by the team captain before class; 4. The teacher gave her grade after providing feedback and making amendments. (The final team grade was decided taking peer/self-grading into consideration when worksheets were collected after class). 	<ol style="list-style-type: none"> 1. Students thought individually or worked with their neighbours for two minutes; 2. A student volunteered his/her answers, or teacher provided answers herself if there was no volunteer; 3. The teacher provided feedback and graded the volunteer’s performance.
II. Introduction of objectives (e.g. how to identify special quantities, make recommendations, and give praise) (15 minutes)	
The intervention group	The comparison group
<ol style="list-style-type: none"> 1. The teacher introduced the objectives to the whole class and explained them by giving brief examples; 2. Students worked in their teams for five minutes to brainstorm more expressions and words to perform the relevant functions; 3. Teamwork was structured and processed following the similar procedures specified above from step (2) to (4). 	<ol style="list-style-type: none"> 1. The teacher gave whole-class instruction on learning objectives and illustrated language functions in detail by providing examples of relevant expressions and words.

(continued)

Appendix 7: Continued

III. Four lead-in questions (15 minutes) 1. <i>Many of my friends keep pets. They think pets bring a lot of fun to their lives. Why do people keep pets?</i> 2. <i>Dogs and cats are commonly kept as pet. Which do you think are better pets, dogs or cats?</i> 3. <i>Pets are usually kept inside. Is it cruel to keep pets inside an apartment all day? Explain why?</i> 4. <i>Some people think man should not eat any kind of meat? Do you think it's OK to eat meat? Do you think it's OK to eat dog meat?</i>	
The intervention group	The comparison group
1. Each question was randomly assigned to three or four teams for discussion; 2. Teamwork was organized and processed following the similar procedures specified in the part of warm-up questions. However, this time four teams instead of one were randomly chosen, one for each question.	1. Each student selected one question that was interesting to them and then worked on it individually or with neighbours; 2. Four volunteers were called upon to present answers for each question. (The teacher provided answers when there was no volunteer).
IV. Listening skills for identifying special expressions of quantities. (Students were required to listen to the following four sections of short conversations and answer questions.) (10 minutes) <i>e.g. A. What are you so happy about?</i> <i>B. Instead of being given an even dozen cookies, we've been given a baker's dozen.</i> <i>Question: Why is the man so happy?</i> <i>A) Twelve cookies were given to them.</i> <i>B) They were given more cookies than they expected.</i> <i>C) The baker was not able to make the cookies.</i> <i>D) The cookies taste much better than the man had hoped.</i>	
The intervention group	The comparison group
1. Students listened to the conversations twice and marked out correct answers; 2. Students worked in their teams for three minutes to compare answers with teammates and discussed their disagreements; 3. Students listened to the conversations again to check through their answers; 4. Students worked out their agreed answers in their teams; 5. Teamwork was processed following the similar procedures specified in the part of warm-up questions.	1. Students listened to the conversations twice and marked out the correct answers; 2. The teacher or volunteers explained answers; 3. Students listened to the conversations again so as to pick up what they missed and achieve a better understanding.

(continued)

V. Introduction of assignment tasks to be checked in the following session: situation dialogue creation and role-play (5 minutes)

Situation 1: Your good neighbour's wife died a couple of weeks ago, and he is feeling lonely. You and your friend are discussing what to buy to cheer him up. Initially, you want to buy a bunch of flowers, but your friend suggests a dog.

Situation 2: You bought a Discman and it has given you a lot of pleasure. Your friend came and saw it, and you highly recommend it.

Task Requirements:

According to the two situations given above, you are required to create dialogues, using relevant expressions and phrases we have just learned in this unit. The minimum length for each dialogue is two minutes.

The intervention group	The comparison group
<ol style="list-style-type: none"> 1. Students created dialogues in their teams after class according to the requirements on the work sheet; 2. Each home team split into two pairs and role-played the dialogue after class; 3. Teamwork was processed following the similar procedures specified in the part of warm-up questions. 	<ol style="list-style-type: none"> 1. Students found themselves a partner, and created dialogues with partners after class; 2. Students practised with their partners after class; 3. Students volunteered to role-play during the next class and their performance was individually assessed by the teacher.

Appendix 8: Contrasts between the two different methods for teaching writing

<p>Task requirements: Choose one from the following three given topics and write a short paragraph with a set of sequential of actions. The paragraph length is no less than 60 words. Attention should be paid to accuracy of spelling, grammar, punctuations, and appropriate use of cohesive devices for linking details to support a general statement.</p> <ol style="list-style-type: none"> 1. He is always willing to help. 2. Her English is good. 3. The weather is getting much warmer. 	
The intervention group	The comparison group
<ol style="list-style-type: none"> 1. Students met in teams and selected one topic to write on; 2. Teammates worked together to generate ideas and brainstorm relevant cohesive devices for writing; 3. Students wrote individually and then exchanged their drafts with teammates for corrections and feedback; 4. Students revised their work accordingly; 5. At the preset assignment submission time (which was usually the following week), some students were selected as team representatives to hand in their revised writing, which carried the grades given by the writers themselves and peer reviewers; 6. The teacher worked out the final grade and provided feedback in the following session. The graded pieces of writing were also circulated among students to welcome additional feedback. 	<ol style="list-style-type: none"> 1. Students individually decided on which topic to write about. 2. Students worked on their writing individually; 3. At the preset assignment submission time, some students were selected to hand in their writing for grading; 4. The teacher gave grades and provided feedback in the following session.

Appendix 9: Contrasts between the two different methods for teaching vocabulary

Type	<p>Requirements: Students were asked to make a list of words relevant to a particular new word being learned. For instance, four tasks were designed to learn the four words: cabbage, slim, suspect and harden.</p> <ol style="list-style-type: none"> 1. List at least ten items in English that are members of the vegetable family. 2. List at least ten adjectives that are used to describe the shape of one's body. 3. List at least five words that can be used to substitute for the word "suspect" in the sentence "I suspect he got it wrong". 4. List at least six words that share the same word formation as "harden", that is, an adjective + "-en" = a transitive verb. 	
One	The intervention group	The comparison group
	<ol style="list-style-type: none"> 1. Students worked in teams and brainstormed the list as quickly as possible; 2. The team that first worked out the list raised their hands; 3. A member from the team volunteered to put the words on the blackboard and gave necessary explanations; 4. The teacher and audience teams gave feedback and the team won bonus points. 	<ol style="list-style-type: none"> 1. Students worked individually and volunteered to list the words on the blackboard; sometimes the teacher listed the words herself if there was no volunteer.
Type	<p>Requirements: Students were asked to create a set of meaningful coherent sentences by using four given words that were selected from the list of new words they had just learned. (If students could not work it out within the given time, the level of difficulty could be reduced by using any three of the four given words.)</p>	
Two	The intervention group	The comparison group
	<ol style="list-style-type: none"> 1. Students worked in teams on the task; 2. The team that first worked out the task raised their hands; 3. A member from the team volunteered to give the sentences. 4. The teacher and audience teams gave feedback and the team won bonus points. 	<ol style="list-style-type: none"> 1. Students worked individually and volunteered answers; usually it was the teacher who provided the first sentence and worked with the whole class to complete the task.

(continued)

Appendix 9: Continued

Type	Requirements: Students were asked to tell the differences between some easily confusing words, phrases, or synonyms. For instance, use examples to show the differences between “result in...”, “result from...”, “as a result (of ...)”. (Difficult vocabulary tasks can be left as assignments so that students can refer to dictionaries, references books, etc).	
	The intervention group	The comparison group
Three	<ol style="list-style-type: none"> 1. Students worked in teams on the task (during or after class); 2. Several students volunteered the answers after teamwork (or in the next class if the task was an assignment); 3. The teacher gave feedback and the teams of the volunteers won bonus points. 	<ol style="list-style-type: none"> 1. Students worked individually and volunteered answers. The teacher provided answers if there was no volunteer.

Appendix 10: A sample of College English Test paper

A sample of written test for first-year non-English majors

考试注意事项:

1. 本考试为闭卷, 要求单独完成, 考试时间 2 小时;
2. 本试卷包括试题册, 机读卡以及答题纸; 16-25 题以及写作部分要求在答题纸上完成, 其余各题对号填涂在机读卡。
3. 本考试提前 30 分钟收取机读卡, 答题纸和试题册考试结束收回。

Part I Listening Comprehension (25%)

Section A

Directions: *In this section, you'll hear 5 short conversations. After each conversation, a question will be asked about what was said. The conversation and question will be read only once. Listen carefully and choose the best answer to each question.*

- | | |
|---|---|
| 1. A. She's careless. | B. She likes George very much. |
| C. She's not concerned with George's health. | D. She doesn't care what George said. |
| 2. A. The speed limit was not clearly marked. | B. The speed limit is 10 miles per hour. |
| C. The speed limit is 40 miles per hour. | D. The speed limit is 50 miles per hour. |
| 3. A. On the grass. | B. Near the pool. |
| C. At home. | D. In the car. |
| 4. A. Philadelphia. | B. California. |
| C. Florida. | D. Arizona. |
| 5. A. Jason Daniels isn't home right now. | B. The caller dialled the wrong number. |
| C. Jason can't come to the phone right now. | D. Jason doesn't want to speak to the caller. |

Section B

Directions: *In this section, you'll hear a long conversation and a passage. The conversation and passage will be read twice. At the end of them, you'll hear some questions. Listen carefully and choose the best answer to each question.*

Questions 6 to 10 are based on the following conversation:

- | | |
|------------------------------------|-------------------------------------|
| 6. A. From one of his friends. | B. From an ad put by the store. |
| C. From the personnel manager. | D. From a clerk in the store. |
| 7. A. Textile technology. | B. Dress designing. |
| C. Business administration. | D. Advertising design. |
| 8. A. To report to his department. | B. To start training. |
| C. To offer more information. | D. To fill out an application form. |
| 9. A. Sporting items. | B. Man's clothing. |
| C. Market analysis. | D. Staff training. |

10. A. Three weeks. B. Three months.
C. Three days. D. Three hours.

Questions 11 to 15 are based on the following passage:

11. A. With the help of his father. B. By applying for a scholarship.
C. By taking a part-time job. D. By getting loans from the bank.
12. A. To have a holiday with his girl friend.
B. To persuade his girl friend to go back to him.
C. To visit one of his friends there.
D. To earn some money for his next year's study.
13. A. He found it hard for him to remain in the top 20% in the class.
B. His grades came down without the help of his girl friend.
C. He found the courses became more and more difficult every year.
D. He felt so tired that he could not keep working hard.
14. A. He promised to send his son more money.
B. He advised his son to forget his girl friend.
C. He showed his disappointment with his son.
D. He didn't believe his son had done his best in his study.
15. A. He accepted his father's help.
B. He transferred to another college.
C. He concentrated on his schoolwork again.
D. He won back his girl friend at last.

Section C

Direction: *There is a passage with 10 blanks in this section. You'll hear the passage read twice. Listen carefully and fill in the blanks with what you have heard.*

After a busy day of work and play, the body needs to rest. Sleep is (16) _____ for good health. During this time, the body (17) _____ from the activities of the previous day. The rest that you get while sleep (18) _____ your body to prepare itself for the next day.

There are four (19) _____ of sleep, each being a little deeper than the one before. As you sleep, your muscles relax little by little. Your heart (20) _____ more slowly, and your brain slows down. After you reach the fourth level, your body (21) _____ back and forth from one level of sleep to the other.

Although your mind slows down, from time to time you will dream. Scientists who study sleep state that when dreaming (22) _____, your eyeballs begin to move more quickly (although your eyelids are closed). This stage of sleep is called REM, which (23) _____ for rapid eye movement.

If you have trouble falling asleep, some people (24) _____ breathing very slowly and very deeply. Other people believe that drinking warm milk will help make you drowsy. There is also an old suggestion that (25) _____ sheep will put you to sleep!

Part II Reading Comprehension (30%)

Directions: *There are 3 passages in this section. Each passage is followed by some questions or*

unfinished statements. For each of them there are four choices marked A, B, C and D. Please choose the most suitable one.

Passage 1

Questions 26 to 30 are based on the following passage.

“Keep an eye on Esther. I’ll be back in a second,” Joy Warren said to her three-year-old son Stephen, who was sitting in the back of the Buick (别克车). She didn’t like leaving the children alone in the car, but the baby was sleeping soundly. And it would only be a moment.

She had hardly walked 40 yards when she saw the car moving. It headed straight towards the river. Unable to swim, Joy shouted, “My babies are in that car!”

Daniel Whitehead, a 17-year-old student, was walking by the river when the Buick crashed into the water just yards ahead. Without thinking, Daniel jumped in. Though a competitive swimmer, he was shocked by the icy chill.

Two minutes earlier, Skip Womack had pulled to a halt as the Buick ran in front of him. Now seeing it hit the water and hearing Joy’s cries, Skip got out of his truck and jumped into the water. He had only one thought: If I don’t get them out, they’ll drown.

Daniel reached the car and grabbed a door handle. But the water was only four inches beneath the window, and the door wouldn’t open. With one powerful punch, Daniel and Skip broke a window. Daniel reached inside and lifted Stephen out. He placed him on his back and set out for shore. At the same time, Skip squeezed himself through the window. He managed to free Esther from beneath her seat belt. After he got out of the car with the baby, he held her over the water and swam toward the shore. All this took place just seconds before the Buick disappeared beneath the water.

Later, Skip thought of his wife and children—how close he’d come to leaving them behind. He thought of the miracle he’d lived through, and how two children were still alive because they happened to be in the right place at the right time.

26. Why did Joy leave her children in the car?
- A. She did not like shopping with a child in hand.
 - B. She did not like waking up her baby.
 - C. Stephen was big enough to take care of his sister.
 - D. It was icy cold outside.
27. Daniel and Skip, who saved the lives of the two kids, _____.
- A. had been good friends
 - B. were two close friends of Joy Warren’s
 - C. were Joy Warren’s neighbours
 - D. were strangers before the accident
28. How did Daniel get Stephen out of danger?
- A. He squeezed into the car and carried him out.
 - B. He pulled him out through the broken window.
 - C. He freed him from his seatbelt before he got him out.
 - D. He held him over the water and swam back to shore.
29. What happened to Joy Warren’s Buick?
- A. It had four inches of water in it.
 - B. One of its windows could not be opened.

Passage 3

Questions 36 to 40 are based on the following passage.

Why is it that the more connected we get, the more disconnected I feel? Every advance in communications technology is a step back from the closeness of human interaction. With e-mail and instant messaging over the Internet, we can now communicate without seeing or talking to one another. With voice mail, you can conduct entire conversations without ever reaching anyone. If my mom has a question, I just leave the answer on her machine.

As almost every imaginable contact between human beings becomes automatic by machine, the alienation quotient goes up. You can't even call a person to get the phone number of another person anywhere. Directory assistance is almost always fully automatic by machine. Pumping gas at the station? Why say good-morning to the worker when you can swipe (刷卡) your credit card at the pump and save yourself the bother of human contact? Placing money at the bank? Why talk to a clerk who might live in the neighborhood when you can just put your credit card into the ATM?

Pretty soon you won't have the hard task of making eye contact at the grocery store. Some grocery chains are using a self-scanner so you can check yourself out, avoiding those annoying clerks who look at you and ask how you are doing.

36. The writer's attitude towards advances in communications technology may be described as _____.

- A. uninterested B. unconcerned C. positive D. critical

37. If his mom has a question, he _____.

- A. will leave her a message through voice mail
B. will email her
C. will try to get her a voice machine
D. will find ways to reach her

38. Judging from the context, the word "alienation" in line 2, paragraph 2 means _____.

- A. relationship B. closeness C. strangeness D. stress

39. With the rapid development of high technology, people don't need to _____.

- A. completely rely on manpower B. say good morning to workers
C. make eye contact with clerks D. see or talk to one another

40. It can be inferred that the writer _____.

- A. is pleased with the modern pace of life
B. doesn't want anyone to bother him
C. feels more separated from others
D. sings the praises of communications technology

Part III Vocabulary and Structure (10%)

Directions: *There are 20 incomplete sentences in this part. For each sentence there are four choices marked A, B, C and D. Choose the answer that best completes the sentence.*

41. My wife is always angry with me; I can hardly _____ her bad temper.

- A. catch up with B. keep up with C. come up with D. put up with

42. Because of the mist, none of those who saw the flying object could tell its _____ shape.
A. precise B. likely C. steady D. rigid
43. The _____ Chinese donated a lot of money to build a primary school in that small village.
A. abroad B. foreign C. alien D. overseas
44. Thirty miles away from the town, the robbers _____ the car and disappeared.
A. approached B. ground C. repaired D. abandoned
45. The little boy went to school merrily, _____ by his parents and friends.
A. accused B. appointed C. accompanied D. accounted
46. After hundreds of experiments, they finally made a major _____ in cancer research.
A. breakthrough B. breakup C. breakdown D. break-in
47. A railway ticket should indicate the place of departure as well as the _____ of the trip.
A. arrival B. image C. destination D. sequence
48. _____ the look on his face, he doesn't quite understand what the speaker is saying.
A. To judge by B. Judging by
C. Judge by D. To be judged by
49. I'll never know all that was in his mind, _____.
A. nor will anyone else B. nor won't anyone else too
C. nor anyone else will D. nor will anyone else either
50. Their average cost of housing _____ 300 Yuan each month.
A. equals to B. reaches to C. adds to D. amounts to
51. "_____ I miss the train?" Mary asked anxiously.
"Don't worry. There's enough time." Jim comforted her with smiles.
A. How about B. What about C. If only D. What if
52. I hurried to the shop only _____ I'd left all my money at home.
A. finding B. found C. founded D. to find
53. I had been self-employed for so long that I couldn't imagine _____ for someone else.
A. to have to work B. working C. to work D. having worked
54. _____ coming of the Space Age, a new area has been added to the study of the planets.
A. While B. It is the C. When the D. With the
55. _____ I admit that the problems are difficult, I don't agree that they can't be solved.
A. For fear that B. Because C. While D. Until
56. With the job _____, he went to see a film with his colleagues.
A. done B. doing C. to do D. do
57. Measures had to be taken in face of the housing problem that _____ in the city.
A. rose B. raised C. arose D. produced
58. _____ what you intended, I should not have wasted my time trying to explain matters to you.
A. When I had realized B. Had I realized
C. Realized had I D. Had realized I
59. There's no limit _____ what you can do if you try.
A. of B. to C. within D. for
60. Will you take a message for me?
"_____."
A. You bet B. I bet C. You are sure D. I'm sure

Part IV Cloze (10%)

Directions: Fill in each of the numbered blanks with the best of the four choices given.

People who use Standard English do not use exactly the same kind of language all the time. At school, people talk more (61) during break and in the cafeteria than they (62) in the classroom or in the office. Secretaries talk more (63) during a coffee break than they do on the (64) when talking to a client. Writing is (65) more formal than speech: What we write down is often a little more important and a little more serious than (66) we just casually mention.

We can say there are roughly (67) main varieties of Standard English. Informal English is the English of conversation and the (68) letter. Formal English is the English used in a public speech, the business letter, and most books and magazines.

Everyone (69) situations (70) which formal English is expected, a letter to public agencies or to a newspaper, a public discussion of a serious (71), a report of an accident, a business memo, letters of (72) for jobs or for college admission, and papers or oral reports in the classroom.

The differences (73) informal and formal English are like the differences between the clothes you wear to the beach and (74) you wear to school. One set of clothes is right for one (75); the (76) set is right for the other. Just so, informal English is right for everyday conversation. A more formal kind of English is (77) speeches or papers (78) serious issues. Through experience, you can learn how to (79) your dress according to occasions. (80), it is through experience that you learn to vary your language and discover which words and phrases are appropriate for formal occasions.

- | | | | |
|--------------------|----------------|-----------------|---------------|
| 61. A. informally | B. formally | C. seriously | D. calmly |
| 62. A. did | B. have | C. do | D. had |
| 63. A. friendly | B. casually | C. seriously | D. formally |
| 64. A. office | B. classroom | C. telephone | D. radio |
| 65. A. generally | B. specially | C. specifically | D. hardly |
| 66. A. that | B. which | C. when | D. what |
| 67. A. three | B. two | C. several | D. more |
| 68. A. personal | B. public | C. business | D. formal |
| 69. A. meets | B. encounters | C. expects | D. applies |
| 70. A. in | B. on | C. for | D. with |
| 71. A. issues | B. speech | C. report | D. issue |
| 72. A. requirement | B. application | C. report | D. speech |
| 73. A. of | B. among | C. between | D. from |
| 74. A. that | B. those | C. which | D. what |
| 75. A. place | B. occasion | C. conversation | D. situation |
| 76. A. other | B. others | C. another | D. either |
| 77. A. to | B. on | C. in | D. for |
| 78. A. on | B. in | C. for | D. of |
| 79. A. differ | B. different | C. vary | D. various |
| 80. A. Likewise | B. Therefore | C. So | D. Contrarily |

Part V. Writing (15%)

Directions: *For this part, you are required to complete a short piece of writing on the topic of **Friendship**. You should write no less than 120 words according to the outlines given below in Chinese.*

1. 友情在生活中很重要。
2. 什么是真正的友情。
3. 应该怎样维持真正的友情。

A sample of spoken English test (20 minutes)

Part 1 (5 minutes)

Examiner: Hello, everyone. I'd like each of you to give a brief introduction about yourself? This includes your name, your major, and some other background information you'd like to tell us. Let's start from you ...

Part 2 (10 minutes)

Examiner: Now that we know each other, let's have a talk on the theme of transport, especially in the city. Each of you is required to give an individual presentation on the topic you are given. Your presentation is limited to one and a half minutes. So please don't worry if I interrupt you when your time is up. Now, let start from student A. Your topic is ...

The four individual questions are (1.5 minutes for each):

Student A: Everyone living in this city must have some experiences of getting stuck in a traffic jam. Please share with us one of your experiences.

Student B: What forms of transport do you often use? why?

Student C: Do you think the use of private cars should be encouraged in China? Give reasons, please.

Student D: Is it a good idea to encourage people to ride bicycles?

Examiner: Right. We have shared some information and ideas about city transport and relevant problems. Now you will be given four minutes for a group discussion. During the discussion you may agree or disagree with each other, or may ask each other for clarification and explanation. Your performance will be judge according to your contributions to the discussion. The topic for group discussion is ...

Topic for group discussion (4 minutes):

What do you think is the most suitable forms of transport in the future for big cities like Beijing and Shanghai? Give reasons and illustrate with examples.

Part 3 (5 minutes)

Examiner: Now I'd like to check with you about some points you mentioned in your discussion. A, I have a question for you. What did you mean by saying ...?

Some other possible questions include:

- 1. I'd like you to give an example to explain ...*
- 2. How did you get the information that ...?*
- 3. Do you agree with ... that...?*
- 4. Your idea is very interesting. But do you think it's feasible to...?*

Appendix 11: Scoring criteria for speaking on the CET

Score range	Accuracy and range of vocabulary and grammar	Length of utterance and discourse management	Flexibility and appropriateness
10-9	<ul style="list-style-type: none"> ● Generally being able to use accurate vocabulary and grammar; ● Being able to communicate through a good range of vocabulary and grammatical structures; ● Having good pronunciation, in spite of some accent which does not affect understanding. 	<ul style="list-style-type: none"> ● Being able to carry on relatively long and coherent conversations on relevant topics, in spite of occasional pauses for being unable to find right words in time. 	<ul style="list-style-type: none"> ● Being able to carry on conversations naturally and actively; ● Being able to use the appropriate language to suit the specific context.
8-7	<ul style="list-style-type: none"> ● Involving a few grammatical and lexical mistakes, which do not affect communication significantly; ● Being able to use a range of vocabulary; ● Having satisfactory pronunciation. 	<ul style="list-style-type: none"> ● Being able to carry on coherent conversations which are short on most occasions; ● Having frequent pauses when searching for the right words or organizing thoughts, which sometimes affect communication; 	<ul style="list-style-type: none"> ● Being able to actively participate in group discussion, but sometimes failing to talk well to the point and interact with others effectively; ● Being able to use the appropriate language to some extent.
6-5	<ul style="list-style-type: none"> ● Involving some grammatical and lexical mistakes, which sometimes affect communication; ● Communicating with only limited vocabulary and simple grammatical structures; ● Having problems with pronunciation, which sometimes affect communication. 	<ul style="list-style-type: none"> ● Only being able to carry on simple and short conversations; ● Having frequent and long pauses when searching for words and organizing thoughts, which affect communication. 	<ul style="list-style-type: none"> ● Being unable to participate in group discussion actively; ● Sometimes being unable to adapt to a new topic.

(continued)

Appendix 11: Continued

Score range	Accuracy and range of vocabulary and grammar	Utterance length and discourse management	Flexibility and appropriateness
4-3	<ul style="list-style-type: none"> ● Involving quite a lot grammatical and lexical mistakes, which have a major effect on understanding; ● Lacking necessary vocabulary and grammatical structures, which has a major effect on communication; ● Having poor pronunciation, which significantly affects communication. 	<ul style="list-style-type: none"> ● Only being able to carry on simple and short conversations, which are often incoherent. 	<ul style="list-style-type: none"> ● Being unable to participate in group discussion.

Appendix 12: The Language Learning Orientations Scale in both Chinese and English versions.

Language Learning Orientations Scale (LLOS)

Directions: The following section contains a number of reasons why one might learn English. Besides each one of the following statements, write the number from the scale which best indicates the degree to which the stated reason corresponds with one of your reasons for learning English. Remember that there are no right or wrong answers, since people have different opinions.

1	2	3	4	5	6	7
Does not correspond at all	Corresponds very little	Corresponds a little	Corresponds moderately	Corresponds a lot	Corresponds almost exactly	Corresponds exactly

I am learning English...

Integrated motivation

1. Because knowing the language is a part of who I am and what I do.
2. Because it will enable me to better understand English speakers' life and culture.
3. Because it is important part of how I define myself.
4. Because it will allow me to participate more freely in the activities of English speakers.
5. Because it will allow me to meet and converse with more and varied people.
6. Because it will allow me to gain good friends more easily among English speakers.
7. Because it is a fundamental part of who I am.
8. Because knowing English is an integral part of my life.

Amotivation

9. I don't know; I cannot come to understand why I am studying English.
10. I cannot see why I study English; I couldn't care less about it.
11. Off hand, I can't think of any good reason for why I study English.
12. Honestly, I don't know why I study English; I truly have the impression that I am wasting my time studying English.

Introjected motivation

13. Because I would feel ashamed if I couldn't speak to English speakers in English.
14. Because I would feel guilty if I didn't know the language.
15. Because it is expected of me.
16. Because I would feel embarrassed or ashamed if I didn't know the language.
17. Because it would make me appear more cultured.
18. Because I would feel bad if I didn't know the language.
19. To prove to myself that I am a good student because I speak English.

Identified motivation

- 20. Because I choose to be the kind of person who can speak English.
- 21. Because I think it's a good idea to know some English.
- 22. Because knowing English may be a gateway to new opportunities.
- 23. Because I think it is good for my personal development.
- 24. Because I choose to be the kind of person who can speak more than one language.
- 25. Because knowing English helps me to achieve goals that are important to me.

External motivation

- 26. Because I want to complete an academic requirement.
- 27. In order to get a more prestigious job later on.
- 28. Because it will give me an edge in competing with others.
- 29. To gain the benefits (e.g., job, money, course credit) which taking English will provide.
- 30. Because I think it will someday be useful in getting a good job.
- 31. To gain the benefits that entrance into the English community will provide me.
- 32. In order to have a better salary later on.

Intrinsic motivation

- 33. For the pleasure I experience when surpassing myself in my English studies.
- 34. For the satisfaction I feel when I am in the process of accomplishing difficult exercises in English.
- 35. For the positive feeling that I experience when using English.
- 36. Because I enjoy the feeling of acquiring knowledge about English-speaking people and their way of life.
- 37. For the satisfied feeling I get in finding out new things.
- 38. Because I love doing it; it's fun.
- 39. Because it's a great feeling to be able to use English.
- 40. For the "high" I feel when hearing the English language spoken.
- 41. For the enjoyment I experience when I grasp a difficult grammatical structure in English.
- 42. For the pleasure I experience in knowing more about the English language.

你为什么学英语？

以下的问题包含了许多你可能学习英语的理由。这些理由可能符合你的情况也可能不符合。认真阅读每一句话，请从以下的 7 个尺寸表中选择最符合自己一项，并把对应号码填写在每一个问题左侧的空白线上。（请注意选择因人而异，没有对错，每句话只能填写一个选项）。

1	2	3	4	5	6	7
根本不 符合	几乎不 符合	有点符合	大致符合	非常符合	几乎完全 符合	完全符合

我正在学英语……

Integrated motivation

1. 因为懂得这种语言在一定程度上能反映“我是谁”和“我做何事”。
2. 因为它将能使我更好地了解英语国家人们的生活和文化。
3. 因为它对我的自我定义是重要的一部分。
4. 因为它能让我更自由的参与说英语的人的活动。
5. 因为它能让我结识更多各式各样的人群并与他们交谈。
6. 因为它让我更容易从说英语的人当中交到好朋友。
7. 因为它对于“我是谁”十分重要。
8. 因为懂得英语是我生活不可或缺的一部分。

Amotivation

9. 我不知道；我无法了解我为什么学英语。
10. 我无法明了我为什么学英语；我不怎么在乎。
11. 我无法立即想出任何一个我为什么学英语的好理由。
12. 说真的，我不知道我为什么学英语；我真的觉得学英语是在浪费时间。

Introjected motivation

13. 因为如果我不能用英语交流，我会感到很羞愧。
14. 因为如果我不懂这种语言我会感到内疚。
15. 因为那是别人对我的期待。
16. 因为如果我不懂这种语言我会觉得尴尬或羞愧。
17. 因为它会让我成为更有修养的人。
18. 因为如果我不懂这个语言我会觉得糟糕。
19. 我可能会说英语来向自己证明我是个好学生。

Identified motivation

20. 因为我选择做一个能说英语的人。
21. 因为我觉得懂些英语是个好主意。
22. 因为懂得英语可能是通往新机会的一扇门。
23. 因为我认为它对我自身发展有益。
24. 因为我选择做个能说超过一种语言的人。
25. 因为懂得英语能帮助我实现对我很重要的目标。

Identified motivation

- 26. 因为我想完成所必修的学分。
- 27. 为了以后能找到更显赫的工作。
- 28. 因为它将有利于我与别人的竞争。
- 29. 为了获得学习英语带来的好处（例如找工作，加薪资，和获得学分）。
- 30. 因为我认为它对于日后求职有好处。
- 31. 为了获得进入英语圈子才能得到的益处。
- 32. 为了以后能获得更好的薪资。

Identified motivation

- 33. 因为我从学英语过程中体会到了超越自我的快乐。
- 34. 因为在完成困难的英语作业的过程中，我感到满足。
- 35. 因为在使用英语的过程中，我有积极向上的感觉。
- 36. 因为我能获得有关英语国家人民及其生活方式的知识，我享受这种获得知识的感觉。
- 37. 为了我发现新事物时那种满足感。
- 38. 因为我乐在其中，英语很有趣。
- 39. 因为使用英语的感觉很好。
- 40. 因为当我听到旁人说英语时感到很兴奋。
- 41. 因为当我掌握了困难的英语文法结构时，我感到很满足。
- 42. 为了掌握更多英语时所感受到的乐趣。

Appendix 13: The Social Skills Scale for Chinese College English

Learners in both English and Chinese versions

Social Skills Scale for Chinese College English learners

Directions: In the following section are 40 statements that indicate a feeling or behavior that may or may not be characteristic or descriptive of you. Read each statement carefully. Decide which response will most accurately reflect your answer and write the corresponding number beside the statement. (Keep in mind that there are no right or wrong answers. Write only one response for each statement.)

1	2	3	4	5	6	7
Not at all like me	Very little like me	A little like me	Moderately like me	Very much like me	Almost exactly like me	Exactly like me

Self-confidence

1. If I work hard at something I will eventually be good at it.
2. Luck decides most things that happen to me.
3. I am doing a good job of English language learning.
4. I feel I am making progress in English language learning.
5. Any failure in a presentation would be devastating to me and keep me from participating again.

Sense of cohesion

6. It is hard for me to stay on task in class.
7. I consider peer support indispensable to my English language learning and success.
8. Working with peers can take me further than working alone.
9. I have a strong sense of belonging when working with peers.
10. I offer teammates support and assistance as much as I can and want them to do their best in English.

Initiative in socialization

11. I'm usually the one who initiates conversation.
12. I usually take the initiative to introduce myself to others.
13. I don't speak until others speak to me.
14. I find it very hard to speak in front of class.
15. I feel stressed and uncomfortable when working with others in English class.

Being positive

16. I tend to encourage and praise when working with others.
17. I offer help to those who can't grasp materials learned.

18. I don't give "put-downs" to others, even if they haven't done a good job.
19. I can always find something to learn from others.
20. I believe in the power of encouragement and praise.

Checking for understanding

21. I often ask for clarification and elaboration to have a better understanding.
22. I feel reluctant to ask for help even when I don't understand the learning material.
23. I often offer clarification and elaboration to get my idea across.
24. If there are too many details mentioned I often summarize the main points to generate a clear and overall picture.
25. I often use examples to make myself understood in English class.

Equal participation and accountability

26. I have a clear picture of my personal role in English group work and participate actively.
27. I tend to keep silent in English class, except when I am called to answer some questions.
28. I am willing to share my knowledge and experience.
29. I interrupt politely someone who speaks too long in group discussion to make sure everyone gets a turn.
30. I know my share of contribution is indispensable to group success.

Acceptance and empathy

31. I can easily figure out what others are thinking.
32. It is hard for me to understand others' perspectives.
33. Few people are as sensitive and understanding as I am.
34. Working with a person who is not like me is unbearable.
35. I believe everybody of different English levels can make a contribution to the completion of group tasks.

Conflict management

36. Whenever in a conflict, I like to hear the other side, negotiate the problems and reach a consensus.
37. When there is a disagreement among us, I withdraw and try to escape from it.
38. I often adapt my ideas and behaviour to the group I work with.
39. I am good at negotiating and interacting with others in the process of resolving conflict.
40. I can always reach constructive solutions to problems and find a way out.

中国大学英语学习者问卷调查表

下面的 40 句话分别表示某种感受或行为,这些表达和特征可能符合你的情况也可能不符合。请从以下的 7 个尺寸表中选择最符合自己一项,并把对应号码填写在每一个问题左侧的空白线上。(请注意选择因人而异,没有对错,每句话只能填写一个选项)。

1	2	3	4	5	6	7
根本不像我	几乎不像我	有点像我	大致像我	非常像我	几乎完全像我	完全像我

自信心

1. 如果我努力做某件事情,我最终会把它做好。
2. 对我而言,大多数事情靠的是运气。
3. 我英语学得不错。
4. 我感到我在英语学习方面有进步。
5. 任何一次课堂发言的失败都会让我一蹶不振,进而阻止我今后的参与。

凝聚感

6. 上课时,我很难注意力集中。
7. 我认为个人英语学习和成功离不开同学的支持。
8. 与独自学习相比,与同学一块学习更能让我受益。
9. 我在与同学一起学习时有很强的归属感。
10. 我向同学提供尽可能的支持和帮助,希望他们在英语学业上有最好的表现。

社交主动性

11. 我通常是主动讲话的那个人。
12. 我常常会自动把自己介绍给别人。
13. 如果别人不跟我说话,我是不会先开口的。
14. 面对全班同学讲话,我觉得对我来说很难。
15. 在英语课上的小组活动中,我感到压力大,不舒服。

积极待人

16. 和别人合作时,我往往给予的是积极的鼓励和赞扬。
17. 我主动帮助那些不能掌握所学知识的同学。
18. 即使别人做得不好,我也不会奚落他们。
19. 我总能从别人身上找到值得自己学习的东西。
20. 我相信鼓励和表扬的力量。

理解核实

21. 为了能更好的理解,我常常请别人澄清观点或详述细节。
22. 即使我不理解所学内容或不能完成作业时,我也不愿意请求别人帮助。
23. 为了把自己的意思讲明白,我常常主动澄清观点或详细解释。

- 24. 如果涉及到太多的细节，我常常总结要点，使别人有一个清晰全面的理解。
- 25. 在英语课上我常常会用具体实例让别人理解我的想法。

平等参与和个体责任

- 26. 在英语小组活动中我清楚自己的个人角色和责任并积极参与。
- 27. 在英语课上我往往保持沉默，除非被点名回答问题。
- 28. 我愿意与别人交流自己的知识和经历。
- 29. 在小组讨论中，如果有人讲话的时间太长，我会礼貌地打断他，以便其他人也有讲话的机会。
- 30. 我明白，个人的贡献是成功的小组活动必不可少的因素。

接受理解他人

- 31. 我很容易能知道别人在想什么。
- 32. 对我而言很难理解别人的观点。
- 33. 很少人能像我一样生性敏感和善解人意。
- 34. 和一个与自己有很大差距的人同处，是难以容忍的。
- 35. 我相信，虽然每个人英语水平不同，但他们能为小组任务的完成做出自己的贡献。

冲突处理

- 36. 无论什么时候在冲突中，我都愿意听取对方意见，协商问题，达成共识。
- 37. 当大家意见相左时，我往往会抽身撤出，想法逃避。
- 38. 我常常调整自己的观点和行为，努力与小组其他成员相协调。
- 39. 在解决冲突过程中，我擅长和别人协商和交流。
- 40. 针对问题，我总能找到建设性的解决办法和出路。

Appendix 14: Between-group comparisons of ages, years of learning English, and pre-test scores in 19 areas on the three measures of the CET, the LLOS, and the SSS-CCEL

Variable	Group	Mean	Standard deviation	Sig. (t-test)
Age	Intervention	19.596	0.846	0.861
	Comparison	19.563	1.070	
Years of learning English	Intervention	8.173	2.130	0.438
	Comparison	7.875	1.645	
Listening	Intervention	14.750	3.458	0.851
	Comparison	14.896	4.274	
Speaking	Intervention	6.906	1.151	0.060
	Comparison	6.527	0.790	
Reading	Intervention	21.577	3.826	0.871
	Comparison	21.708	4.267	
Writing	Intervention	8.346	2.132	0.448
	Comparison	8.000	2.415	
Vocabulary	Intervention	14.106	2.042	0.107
	Comparison	13.344	2.632	
Intrinsic motivation	Intervention	36.539	9.022	0.991
	Comparison	36.563	11.616	
Integrated motivation	Intervention	28.846	5.539	0.911
	Comparison	28.667	9.762	
Identified motivation	Intervention	27.539	5.717	0.884
	Comparison	27.354	6.887	
Introjected motivation	Intervention	22.154	6.102	0.917
	Comparison	22.292	7.074	
External motivation	Intervention	33.154	6.044	0.247
	Comparison	31.604	7.249	
Amotivation	Intervention	23.039	4.121	0.812
	Comparison	22.833	4.459	
Self-confidence	Intervention	21.577	3.226	0.530
	Comparison	22.000	3.495	
Sense of cohesion	Intervention	17.885	3.059	0.530
	Comparison	17.458	3.701	
Initiative socialization	Intervention	21.269	3.768	0.924
	Comparison	21.188	4.771	

(continued)

Appendix 14: continued

Variable	Group	Mean	Standard deviation	Sig. (t-test)
Being positive	Intervention	21.385	3.810	0.846
	Comparison	21.208	5.095	
Checking for understanding	Intervention	21.808	4.215	0.552
	Comparison	22.333	4.582	
Equal participation and accountability	Intervention	18.404	2.614	0.585
	Comparison	18.771	3.985	
Acceptance and empathy	Intervention	21.269	2.794	0.054
	Comparison	22.708	4.332	
Conflict management	Intervention	20.731	3.861	0.325
	Comparison	21.479	4.141	