1-20-2006

Effects of Cooperative Learning on Motivation, Learning Strategy Utilization, and Grammar Achievement of English Language Learners in Taiwan

Hui-Chuan Liao
University of New Orleans

EFFECTS OF COOPERATIVE LEARNING ON MOTIVATION, LEARNING STRATEGY UTILIZATION, AND GRAMMAR ACHIEVEMENT OF ENGLISH LANGUAGE LEARNERS IN TAIWAN

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
The Department of Curriculum and Instruction

by

Hui-Chuan Liao

M.Ed., University of New Orleans, 1999

December 2005

To my husband and soul mate, Tien, who gives me love and support.
To my son and best friend, Jesse, who gives me hope and joy.
To my parents, who give me life and everything.

In loving memory of my mom,
who died of pancreatic cancer during the first year of my doctorate.
Mom's fortitude in fighting the illness
taught me to remain strong and determined
in completing this degree amid the aftermath of Hurricane Katrina.

ACKNOWLEDGEMENTS

I would like to acknowledge the Graduate School and the Department of Curriculum and Instruction at the University of New Orleans for awarding me the Crescent City Doctoral Scholarship and the Graduate Assistantship throughout my doctoral study. Without the awards, the completion of the doctorate would have been financially impossible. Sincere appreciation also goes to my dedicated committee members who provided immense guidance and assistance throughout my Ph.D. program, especially after Hurricane Katrina devastated New Orleans and caused severe damage to their homes.

I was blessed with a knowledgeable, thorough, and supportive major professor and committee co-chair, Research Professor John Barnitz. I am grateful for his guidance at every step of my doctorate. I thank him for allowing me to pursue my interests, trusting in my ability to make major decisions, constantly challenging me with questions that enriched my learning experience, and repeatedly encouraging me with kind words. Even though he was displaced and struggling in the recovery process after the hurricane, he still rendered me timely support as if guiding my dissertation was his number one priority. Without his quick response each time I turned in my draft, I could not have finished the dissertation in time to graduate this semester.

Heartfelt gratitude goes to my methodologist and committee co-chair, Dr. Jeffrey Oescher, for believing in my potential for pursuing the doctorate when I was only a first-year graduate student. I thank him for working closely with me on research design,

instrument construction and evaluation, statistical analysis, and interpretation of data. He impressed me by demonstrating how enthusiastic a professor could be about a student's learning and research. I thank him for making me understand how profoundly a professor could influence a student's life forever. When I realized that I could miss the defense deadline, I became panicked. I thank him for calming me down by assuring me, "I promise I'll get you finished if I do nothing else this semester!"

A debt of gratitude is owed to Professor Patricia Austin, who took me under her wings like a big sister and called me a friend. I am grateful for her valuable input in the theoretical and methodological aspects of my dissertation. I also thank her for editing my dissertation with such meticulousness. Throughout the program and especially after the hurricane when her house had been ruined, the professional assistance and emotional support she continued to offer was simply incredible.

My gratefulness also goes to Professor Renee Casbergue, Chair of Curriculum and Instruction, for showing me how to critique literature analytically and guiding me through the process. As her graduate assistant, I witnessed her exceptional administrative abilities, and I know what I have learned from her will forever impact me. To Dr. Richard Speaker, I offer thanks for his insightful and inspiring comments on my dissertation. I appreciate that he always held the standards high and expected me to reach them.

A special thank you goes out to Professor Robert Slavin at the Johns Hopkins University and Professor Merrill Swain at the University of Toronto for generously sharing their expertise in cooperative learning and second language acquisition. My gratitude extends to Professor Elaine Horwitz at the University of Texas at Austin for offering valuable recommendations on my dissertation when we met at the TESOL 2005

Convention. My sincere appreciation goes to Professor Wilbert McKeachie at the University of Michigan for allowing me to use and modify the MSLQ questionnaire for the present study. Special appreciation goes out to Dr. Yi-Guang Lin at the University of Michigan for reviewing my Chinese translation of the MSLQ, giving me so many insightful suggestions on refining my research design and instrument, and leading me to new directions for future research.

My earnest appreciation goes to members of the Cherne family. I thank Joaan and Bill for providing immense support throughout my doctorate. I thank the couple for always making sure I had enough groceries and household supplies and for going to the stores when they asked me to stay at my desk studying. I thank Bill for voluntarily maintaining my car as if that were his responsibility. I thank them for allowing me to tag along during the evacuation and aftermath of Hurricane Katrina. I thank Joaan for constantly telling me that my dissertation was the number one concern and for "grounding me" in my room while she fixed delicious hot meals for us. My heartfelt gratitude goes to Aunt Trisha and Uncle Andy for providing a roof to their nephew's tag-along friend when the sky was falling down on her in the weeks following the hurricane. I thank Aunt Trisha for being there for me, calling up doctors and labs, and handling complicated insurance issues when I was challenged with medical problems. I thank the whole Cherne family for sustaining me when the totality of the dissertation, the hurricane, and the medical issues became almost too much to bear.

Sincere gratitude goes to Frances for being my first statistics teacher and to Ying for her friendship and constant advice. To Michelle, Amber, Hui-Yun, and a friend who chose to remain anonymous, I offer thanks for their assistance with data collection.

Although I was thousands of miles away from my home in Taiwan, the love and support from my family was never discounted by the distance over the Pacific Ocean. Heartfelt thanks go to my son and best friend, Jesse, who gave me the drive to stay focused on my study. I thank him for telling me, “Don’t worry, Mommy. I’ll be a very good boy so that you can finish your study sooner.” His understanding, support, and cooperation so touched my heart that I was determined to invest every possible minute to this dissertation. I have been blessed with such a wonderful child since the day he was born.

I can never thank my parents enough for their boundless love and support, not just during the doctorate but also throughout my life. Dad, thanks for believing in me, always encouraging me to pursue my dream, and telling me that I could achieve anything I strive for. Mom, I am indebted that you were worrying about my study rather than your health the day you were diagnosed with advanced cancer. Although it still hurts, I want to thank you for telling me that you were hanging in there in order to see me in the doctoral robe. You did not get to see me all the way through but I know you would be really proud. I thank you for showing me the capacity of love and always being there for me.

Finally, my deepest gratitude goes to my husband and soul mate, Tien-Szu Pan, Ph.D., for his endless love and support in the past 20 years. He owns an exceptional ability to bring sunshine to my heart when the sky is gray. His wisdom, hardworking spirit, and dedication to education inspire me. I thank him for encouraging me to pursue this degree. I am grateful that he took such good care of our child and raised him to be a young man of such character that I was able to concentrate on my study. He sustained me throughout the doctorate. I could not have completed the degree without him.

TABLE OF CONTENTS

LIST OF TABLES xv

LIST OF FIGURES xviii

LIST OF ABBREVIATIONS xix

ABSTRACT xx

CHAPTER ONE: INTRODUCTION 1

Theoretical Framework 1
Purpose of the Study 3
Research Questions 4
Significance of the Study 5
Definition of Terms 6
Organization of the Study 12

CHAPTER TWO: REVIEW OF LITERATURE 14

Theoretical Underpinnings of Cooperative Learning 14
Motivational Theories 14
Expectancy-Value Theories 15
Attribution Theory 15
Self-Efficacy Theory 16
Self-Worth Theory 17
Model of Task Values 18
Goal Setting Theory 19
Self-Determination Theory 20
Summary on Motivational Theories 23
Social Cognitive Theories 25
Vygotsky 26
Piaget 27
Dewey 29
Bruner 31
Bandura 31
Summary on Social Cognitive Theories 33
The Essence of Cooperative Learning 34
Heterogeneous Grouping 36
Positive Interdependence 36
Outcome Interdependence 36

Process Interdependence....37
Individual Accountability....39
Prior Training of Group Skills....40
Group Processing....41
Equal Opportunities for Success....41
Summary on the Essence of Cooperative Learning....42
Differences between Cooperative Learning and Collaborative Learning....43
Brief History of Cooperative Learning....46
Research on Cooperative Learning in the United States....49
Research on Cooperative Learning in Second Language Instruction....54
Need for Innovation in Taiwan's EFL Education....58
Research on Cooperative Learning in Taiwan's EFL Classrooms....61
Studies that Employed Interviews....62
Studies that Employed Questionnaires....62
Study that Employed Observations....63
Studies that Employed Paper-and-Pencil Tests as Instrument....64
Summary of the Studies by Data Collection Methods....66
Review of Studies by Types of Cooperative Methods and Subject Characteristics....66
Second Language Acquisition and Second Language Learning....67
Theories Emphasizing Natural Input....68
Acquisition-Learning Distinction Hypothesis....68
Criticisms....69
Input (Comprehension) Hypothesis....69
Criticisms....70
Studies that Challenge the Input Hypothesis....73
Theories Emphasizing Consciousness-Raising....74
Attention and Second Language Acquisition....74
Noticing Hypothesis....74
Automaticity....75
Research on Explicit Grammar Instruction for L2 Learners....76
An Integrated Approach....77
Pedagogical Practices for Grammar Teaching....79
Input-Oriented Strategies....79
Explicit Grammar Rule Articulation....79
Promoting Noticing through Exposure....80
Output-Oriented Strategies....80
Sentence Manipulation....80
Modeling Language after Authentic Literary Texts....81
Output-Input Combination....81
Collaborative Dialogue....81
Dialogue Journal....82
Writer Conference....83
The Garden Path Technique....83
Summary on Pedagogical Practices for Grammar Teaching....84
Chapter Conclusion....85

CHAPTER THREE: METHODOLOGY 87

Subjects 89
Sampling Method and Procedure 89
Sample Size 90
Sample Characteristics 90
External Validity Issues 91
Instrumentation 92
Motivated Strategies for Learning Questionnaire 92
Motivation 93
Value 94
Expectancy 94
Test Anxiety 94
Learning Strategies 94
Cognitive and Metacognitive Strategies 95
Resource Management Strategies 95
Technical Characteristics 96
Validity 96
Reliability 97
Questionnaire for the Present Study 98
Selection of Scales 98
Modification and Translation 99
Administration 101
Scoring Procedures and Score Interpretation 102
Reliability 104
Grammar Pretest 104
Validity 105
Reliability 105
Administration 105
Scoring Procedures and Score Interpretation 105
Grammar Posttest 106
Validity 106
Item Selection 110
Pilot Test and Item Analysis 115
Item Evaluation through Cross-Examining the Results of the Item Analysis 117
Administration 120
Reliability 120
Scoring Procedures and Score Interpretation 120
Procedures 121
Research Design 121
Teaching Material 122
Independent Variable 124
Procedure for the Control Group 124
Procedure for the Experimental Group 125

Fidelity of Treatment 128
Dependent Variables 129
Relationships between Variables 129
Internal Validity of the Study 129
History 130
Maturation 130
Statistical Regression 130
Selection 130
Mortality 131
Diffusion of Treatment 131
Compensatory Equalization 131
Compensatory Rivalry 131
Pretesting 131
Instrumentation 132
Data Collection 132
Data Analyses 132
Descriptive Analyses 132
Inferential Analyses 132
Tests for Assumptions Underlying ANCOVA 134
Tests for Assumptions Underlying MANCOVA 135
Exploratory Analyses 136
Chapter Summary 137

CHAPTER FOUR: RESULTS 139

Descriptive Analyses 140
The Sample 141
MSLQ Questionnaires 142
Motivation Scales 142
Self-Efficacy Pretest 142
Self-Efficacy Posttest 143
Task Value Pretest 144
Task Value Posttest 144
Use of Learning Strategies Scales 144
Elaboration Pretest 144
Elaboration Posttest 145
Peer Learning Pretest 145
Peer Learning Posttest 146
Summary on the MSLQ Descriptive Analyses 146
Grammar Tests 147
Pretest 148
Posttest 148
Inferential Analyses 148
Research Question One 149
Self-Efficacy 150
Task Value 151

Research Question Two 152
Elaboration 153
Peer Learning 154
Research Question Three 155
Summary on the Inferential Analyses 156
Exploratory Analyses 157
Exploratory Question A 158
Self-Efficacy 159
Task Value 161
Exploratory Question B 163
Elaboration 163
Peer Learning 164
Exploratory Question C 168
Exploratory Question D 170
Exploratory Question E 173
Chapter Summary 174

CHAPTER FIVE: DISCUSSION 177

Effects of Cooperative Learning on Motivation 179
Effects of Cooperative Learning on Utilization of Learning Strategies 183
Effects of Cooperative Learning on Grammar Achievement 187
Interaction Effects of Instruction and Prior English Level 193
Effects of Cooperative Learning on Performance at Different Cognitive Levels 195
Relations among Grammar Achievement, Motivation, Strategy Use, and More 196
Implications of the Present Study 197
Theoretical Implications 197
Methodological Implications 198
Pedagogical Implications 199
Delimitations and Limitations 202
Delimitations 202
Limitations 203
Suggestions for Further Research 204
Conclusion 208

REFERENCES 212

APPENDIXES 231

A. Confirmatory Factor Analyses with the Lambda-ksi Estimates on the Motivation Items of the MSLQ 232
B. Confirmatory Factor Analyses with the Lambda-ksi Estimates on the Learning Strategies Items of the MSLQ 235
C. MSLQ Pretest Questionnaire 239

D. MSLQ Posttest Questionnaire....242
E. MSLQ Pretest Questionnaire, Chinese Version....245
F. MSLQ Posttest Questionnaire, Chinese Version....247
G. Permission for Use and Modification of the MSLQ Questionnaire....249
H. Item Analysis on the Pilot Test for the Grammar Posttest....251
I. Comparison Chart for the Semester Grade Rubrics of the Experimental and the Control Groups....254
J. Consent Form....256
K. Consent Form, Chinese Version....259
L. Use of Human Subjects Approval....262

VITA....266

LIST OF TABLES

Table 1. Effects of Self-Efficacy....17

Table 2. Summary of Expectancy-Value Theories....19

Table 3. Factors Influencing Goal Setting and Goal Commitment....20

Table 4. Motivated Behaviors by the Degree of Internalization....22

Table 5. Ways to Structure Positive Interdependence within a Cooperative Group....36

Table 6. Possible Role Assignments, Responsibilities, and Gambits....38

Table 7. Key Emphases of Major Cooperative Learning Methods....43

Table 8. Differences between Cooperative Learning and Collaborative Learning....45

Table 9. Breakdown of Effect Sizes by Cooperative Methods....52

Table 10. Breakdown of Effect Sizes by Characteristics of Cooperative Methods....52

Table 11. A Quick View of Pedagogical Practices for Grammar Teaching Reviewed....84

Table 12. Structure of the Motivated Strategies for Learning Questionnaire (MSLQ)....93

Table 13. Reliability Level for Each MSLQ Scale....97

Table 14. MSLQ Questionnaire Items Selected for the Study....100

Table 15. Sample Item That Is Positively Worded....102

Table 16. Sample Item That Is Reverse Coded....102

Table 17. MSLQ Score Interpretation....103

Table 18. Reliabilities of the Four Test Banks....105

Table 19. Sample Table of Contents from the Textbook 107

Table 20. Table of Specifications for the Grammar Achievement Test Bank 109

Table 21. Number of Test Items on Each Level of Cognitive Processes: A Comparison Chart 111

Table 22. Table of Specifications for the Grammar Achievement Pilot Study 114

Table 23. Excerpt of the Item Analysis on the Pilot Test for the Grammar Posttest 117

Table 24. Weekly Themes of the Grammar Lessons for Both the Experimental and Control Groups 124

Table 25. Conversion Table for IIS and IIP 127

Table 26. Criterion of Group Award 127

Table 27. Demographic Information of the Sample 141

Table 28. Mean Scores for the MSLQ Scales 143

Table 29. Mean Scores for the Grammar Tests 147

Table 30. One-Way ANCOVA on Self-Efficacy Posttest Scores with Self-Efficacy Pretest Scores as Covariate 150

Table 31. One-Way ANCOVA on Task Value Posttest Scores with Task Value Pretest Scores as Covariate 151

Table 32. One-Way ANCOVA on Elaboration Posttest Scores with Elaboration Pretest Scores as Covariate 153

Table 33. One-Way ANCOVA on Peer Learning Posttest Scores with Peer Learning Pretest Scores as Covariate 154

Table 34. One-Way ANCOVA on Grammar Posttest Scores with Grammar Pretest Scores as Covariate 155

Table 35. Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Self-Efficacy 159

Table 36. Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Task Value 161

Table 37. Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Elaboration....164

Table 38. Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Peer Learning....165

Table 39. Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Grammar Achievement....168

Table 40. ANCOVAs on "Evaluate" and "Create" Level Posttest Scores with Grammar Pretest Scores as Covariate....172

Table 41. Correlations among English Level, Gender, Motivation, Strategy Use, and Achievement....173

Table 42. Summary of Findings....175

LIST OF FIGURES

Figure 1. Graphic representation of the self-determination theory 21

Figure 2. Slavin's model of cooperative learning 25

Figure 3. Model of triadic reciprocality 32

Figure 4. Quasi-experimental pretest-posttest comparison group research design 121

Figure 5. Group average improvement point (GAIP) worksheet 127

Figure 6. Interaction plot for ranking and instruction on self-efficacy 160

Figure 7. Interaction plot for ranking and instruction on task value 162

Figure 8. Interaction plot for ranking and instruction on elaboration 165

Figure 9. Interaction plot for GEPT level and instruction on peer learning.....166

Figure 10. Interaction plot for ranking and instruction on peer learning.....167

Figure 11. Interaction plot for ranking and instruction on grammar achievement.....170

Figure 12. Possible process of how cooperative learning facilitates elaboration skills for self-study.....185

LIST OF ABBREVIATIONS

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CDA	Critical Discourse Analysis
CIRC	Cooperative Integrated Reading and Composition
CL	Cooperative Learning
EFL	English as a Foreign Language
ESL	English as a Second Language
FLCAS	Foreign Language Classroom Anxiety Scales
GAIP	Group Average Improvement Point
GEPT	General English Proficiency Test
IELP	Intensive English Language Program
IIP	Individual Improvement Point
IIS	Individual Improvement Score
ITS	Individual Test Score
L1	First Language
L2	Second Language
LT	Learning Together
MANCOVA	Multivariate Analysis of Covariance
MSLQ	Motivated Strategies for Learning Questionnaire
SLA	Second Language Acquisition
STAD	Student Teams Achievement Divisions
TAI	Team-Assisted Instruction
TEFL	Teaching English as a Foreign Language
TESOL	Teaching English to Speakers of Other Languages
TGT	Teams-Games-Tournament
TOEFL	Test of English as a Foreign Language
TPR	Total Physical Response
WC	Whole-Class
ZPD	Zone of Proximal Development

ABSTRACT

To examine the effects of cooperative learning on EFL students in Taiwan, a 12-week quasi-experimental pretest-posttest comparison group research study was designed. Two college classes (42 students each) in Taiwan participated in the study, one receiving grammar instruction through cooperative learning and the other through whole-class teaching. Three specific research questions guided the study. The first looked at effects of cooperative learning on motivation, the second on out-of-class strategy use, and the third on grammar achievement. Additional exploratory questions examined these results across subgroups within each class as well as the relationships between the dependent variables. Data were collected via learners' pretest and posttest scores on the dependent variables. The data were analyzed with MANCOVAs, one- and two-way ANCOVAs, simple effects, and Pearson correlations.

Cooperative learning was found to have large positive effects on motivation and strategy use, and medium-to-large positive effects on grammar achievement. Overall, the findings indicated a consistent pattern in favor of cooperative learning over whole-class instruction in teaching the Taiwanese learners English grammar. The results of the exploratory questions indicated that cooperative learning facilitated motivation and strategy use of learners across all subgroups, but more so with those performing at higher and lower levels. Grammar achievement of learners at higher and lower levels was affected positively. Additional analyses also indicated cooperative learning positively

affected learning at higher cognitive levels. Implications for future research and for curriculum and instruction are addressed.

CHAPTER ONE

INTRODUCTION

This study investigates the effects of cooperative learning in an English as a Foreign Language (EFL) setting in Taiwan. The study provides a systematic pretest-posttest comparison through achievement and motivational measures on the effects of cooperative learning with whole-class instruction. The effectiveness of cooperative learning has been supported by a large body of research across different grade levels and subject areas in countries such as the United States, Israel, Lebanon, the Netherlands, and Nigeria (e.g., Abrami, Lou, Chambers, Poulsen, & Spence, 2000; Calderon, Hertz-Lazarowitz, & Slavin, 1998; Ghaith, 2003a, 2003b; Johnson & Johnson, 1989; Slavin, 1995; Vaughan, 2002). After being implemented in American classrooms for over a century, this pedagogy has begun to gain attention and interest from EFL teachers in Taiwan, where EFL instruction is still mostly whole-class, teacher-centered rote grammar-translation and often fails to motivate learning (Babcock, 1993; Lai, 2001; Su, 2003; Yu, 1993). Even though numerous efforts have been made to examine the effects of cooperative learning on Taiwanese EFL learners, overall this pedagogy is still under-researched.

Theoretical Framework

According to Slavin's model of cooperative learning (1995), cooperative learning ultimately results in gains in learning because the process of cooperation prompts motivation and consequential cognitive activities. This model is supported by two major

categories of theories: motivational theories and social cognitive theories, e.g., Bandura's self-efficacy theory (1993), Covington's self-worth theory (1992), Vygotsky's zone of proximal development (1978), and Piaget's social transmission theory (1964).

While competition sometimes sets up a stage for students to strive for success over the failure of others, cooperative learning aims to create an arena for team members to have high expectancy of each other. If a teacher implements a cooperative learning method correctly, every learner, including low achievers and high achievers, is expected to be respected and cherished by their peers. More specifically, higher achievers are valued for their knowledge as well as their ability and willingness to share what they know; low achievers are accepted and also respected for who they are and their willingness to make improvement. Students realize that their group members want them to learn and thrive. They become enthusiastic in helping and encouraging one another to learn (Slavin, 1995).

Based on Slavin's (1995) cooperative learning model, when students have the motivation to learn and to encourage and help one another, a stage is created for cognitive development. Vygotsky (1978) argued that cooperation promotes learning because the process enables learners to operate within one another's "zone of proximal development" (p. 86). Working with peers is academically beneficial because, when learners are closer to one another in their levels of proximal development, they are able to describe things to one another in a simpler way that is easier to be comprehended than being explained by a person with a very different mental stage. Likewise, Dewey (1963) stressed the importance of "active cooperation" in the process of constructing knowledge (p. 67).

Slavin's (1995) cooperative learning model is also supported by cognitive elaboration theories. Walling (1987) maintained that discussion of the subject matter by group members during the process of peer work helps students verbalize and elaborate their initial, immature thoughts. In the process of elaboration, a student apparently has chances to develop ideas from vague to concrete and from preliminary to sophisticated, which might not happen if a student just listens to a lecture and passively receives information. In addition, elaboration leads to active processing of information, cognitive restructuring, and reprocessing of ideas. These in turn can increase practice and help a student learn better and retain information longer than those working alone (Snowman & Biehler, 2005; Dansereau, 1988). In a recent empirical study (Veenman, Denessen, Van Den Akker, & Van Der Rijt, 2005), elaboration was found to be positively related to student achievement.

Slavin's model of cooperative learning has been supported by cognitive and motivational theories as well as by a vast number of experimental studies in the United States and other countries. Nevertheless, the use of cooperative learning in Taiwanese EFL classrooms is still under-researched.

Purpose of the Study

After conducting a review of literature on Taiwanese EFL instruction (detailed in Chapter Two), the researcher realized a need to improve EFL education in Taiwan. Although several studies in Taiwanese settings have shed some light on the potential benefits of cooperative learning, the number of studies on the topic is still limited. The current study, therefore, endeavored to investigate how cooperative learning differs from the whole-class method on both cognitive and motivational measures.

The purpose of the study was to investigate the effectiveness of Slavin's model of cooperative learning (1995, p. 45) using a quasi-experimental research design in the context of an EFL course being taken by college students in Taiwan. The study (1) compared the group receiving cooperative learning (CL, the experimental group) and the group receiving whole-class instruction (WC, the control group) in terms of grammar achievement in an EFL college class in Taiwan; and (2) used the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) to compare and understand the differences of students' motivation and out-of-class use of learning strategies between the experimental and control groups. There are many ways cooperative learning can be implemented. The specific cooperative learning method used in this study was Student Teams Achievement Divisions (STAD) developed by Slavin, which emphasizes group goals, individual accountability, and equal opportunities for success. The specific procedures of STAD will be described in Chapter Three.

Research Questions

The study first attempted to answer the following main research questions:

1. How does motivation differ between the group receiving cooperative learning and the group receiving whole-class instruction?
2. How does utilization of learning strategies beyond class settings differ between the group receiving cooperative learning and the group receiving whole-class instruction?
3. How does grammar achievement differ between the group receiving cooperative learning and the group receiving whole-class instruction?

As a result of the analyses on the above three major research questions, the following exploratory questions were investigated:

A. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on motivation? If so, what is the cause of the interaction?
B. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on out-of-class utilization of learning strategies? If so, what is the cause of the interaction?
C. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on grammar achievement? If so, what is the cause of the interaction?
D. How does student achievement differ between the cooperative learning group and the whole-class instruction group in terms of different cognitive levels?
E. What are the relationships among prior English level, gender, grammar achievement, task value, self-efficacy, use of elaboration strategies, and out-of-class peer collaboration behaviors?

Significance of the Study

Although there is a large body of literature on the positive effects of cooperative learning for native speakers of English, there is a gap in the literature regarding the implementation of peer-tutoring cooperative methods on the college level in Taiwan. Do the benefits of cooperative learning that are generally found on native English speakers also apply to EFL students in a Taiwanese college EFL class? Could cooperative learning provide a solution to Taiwanese EFL educators who are striving to figure out a way to help students who are bored and struggling in a whole-class setting? This study

contributes to the body of knowledge by providing evidence that is needed to verify the existing studies so that Taiwanese EFL teachers can justifiably decide whether to use cooperative learning in their classrooms.

Definition of Terms

For the purpose of the study the following terms have been defined:

Accommodator

An accommodator is a learner who likes to do rather than just think. He or she likes hands-on experience, active learning, risk taking, carrying out solutions, and sharing information with others (Kolb, 1984, 1999).

Achievement Test

An achievement test aims to assess what knowledge and skills students have learned from a particular course or set of materials. An achievement test is usually directly anchored in course objectives. It contrasts with a proficiency test, which aims to assess learners' general ability (Brown, 1996; Nunan, 1999).

Assimilator

An assimilator is a learner who enjoys abstract ideas and creating conceptual models. He or she likes to design experiments, consider alternative resolutions, read, reflect, and analyze quantitative information (Kolb, 1984, 1999).

Collaborative Learning

Collaborative learning is the instructional employment of small groups in which student autonomy is emphasized. Due to the autonomous nature of the approach, the teacher set the goal but does not specify the procedure for the groups to meet the goal. Students choose task roles and decide among themselves how things should be done.

Teacher observation and intervention are minimized to sustain students' self-governance (Adams, 2000; Bruffee, 1999).

Comprehensible Input

Comprehensible input is messages addressed to a learner that, while they may contain grammar and structure that are beyond the learner's current ability level, are made comprehensible by the context in which they are uttered (Krashen, 1985, 2002, 2003).

Co-op Co-op

Co-op Co-op is a cooperative learning method that allows learners to investigate in depth topics they find especially of interest. Co-op Co-op consists of 10 major steps: student-centered class discussion, learning team selection, team building, team topic selection, mini-topic selection, mini-topic preparation, mini-topic presentation, team presentation preparation, team presentation, and, finally, class, instructor, and group evaluations of team and individual performances (Kagan, 1985).

Cooperative Integrated Reading and Composition (CIRC)

CIRC is a comprehensive cooperative program for teaching reading and writing. Pairs of different reading levels work on a series of cognitively engaging activities, including partner reading, story grammar, story retell, writing responses, and spelling practice (Slavin, 1995).

Cooperative Learning (CL)

Cooperative learning is the instructional employment of a particular set of "carefully structured" small group activities that are prescribed by the teacher. In cooperative learning, heterogeneous grouping, positive interdependence, and individual

accountability are emphasized. Within a cooperative learning group, students work together and are formally accountable for their own and one another's learning. Teacher observation and intervention are important in cooperative learning (Adams, 2000; Bruffee, 1999; Oxford, 1997). For the purpose of the current study, the specific cooperative learning method was Student Teams Achievement Divisions (STAD). For more discussion on specific procedures of STAD, see "Research Design" in Chapter Three.

Coverger

A coverger is a learner who likes finding practical uses for ideas and theories, assessing consequences and selecting resolutions, following detailed sequential steps, application-oriented activities, and being given clear goals with a logical sequence to activities (Kolb, 1984, 1999).

Diverger

A diverger is a learner who likes inventive, pioneering activities. He or she likes to create a wide array of ideas, identify problems, collect information, and be personally involved in the learning experience and group activities (Kolb, 1984, 1999).

English as a Foreign Language (EFL)

EFL is the teaching and learning of English to non-native speakers of English in communities where the language is not commonly used for communication by the population at large.

English as a Second Language (ESL)

ESL is the teaching and learning of English to non-native speakers of English in communities where the language is commonly used for communication by the population at large.

Field Dependent

A field dependent is someone who has a tendency of being warm, kind, sensitive, and diplomatic. He or she tends to maintains effective interpersonal relationships (Crozier, 1997).

Field Independent

A field independent is someone who has a tendency of separating feelings and cognitions. He or she tends to be demanding, manipulative, and independent of other people (Crozier, 1997).

Grammar Translation

Grammar translation is a language teaching method based on grammatical analysis and the translation of sentences and texts to and from the learners' first and target languages.

Jigsaw

Jigsaw is an instructional method that aims to help students cooperatively learn new material using a group learning approach. Students are accountable for becoming an "expert" on one part of a lesson and then teaching it to other members of their group (Aronson, Blaney, Stephan, Sikes, & Snapp, 1978; Clarke, 1999).

Jigsaw II

Jigsaw II is a method that combines Jigsaw and Student Teams Achievement Divisions (STAD). Students first go through the procedure as they would in a Jigsaw method. Then scoring and group recognition are based on improvement as in STAD (Slavin, 1995).

Learning Together (LT)

Learning Together is an umbrella term for cooperative learning methods and strategies developed by Johnson and Johnson and their colleagues that emphasize face-to-face interaction, positive interdependence, individual accountability, and group skills (Johnson, Johnson, & Holubec, 1994; Johnson, Johnson, Holubec, & Roy, 1984). One of the most widely used Learning Together methods is the Controversy method (see Johnson & Johnson, 2003, for a list and procedures of the Learning Together methods and strategies).

Peer Learning

Peer learning is an umbrella term for cooperative learning and collaborative learning.

Proficiency Test

A proficiency test aims to assess learners' general knowledge or skills. It contrasts with an achievement test, which aims to assess what students have learned from a particular course or set of materials (Brown, 1996; Nunan, 1999).

Restructuring

Restructuring is changes or reorganization of one's grammatical knowledge.

Rote Learning

Rote learning means learning through repetition with minimal attention to meaning.

Second Language Acquisition (SLA)

This is the common name given to the field. It refers to the learning of another language after the first language has been learned.

Structural Syllabus

A structural syllabus consists of a list of grammatical items, usually arranged in the order in which they are to be taught.

Student Teams Achievement Divisions (STAD)

STAD is a cooperative learning method that emphasizes equal opportunities for success by focusing on students' improvement. The method consists of five main steps: teacher presentation, teamwork, individual quizzes, individual improvement scores, and group recognition (Slavin, 1995). For more discussion on specific procedures of STAD, see "Research Design" in Chapter Three.

Target Language

A target language is the language being learned.

Team-Assisted Instruction (TAI)

TAI is a cooperative learning method designed to teach mathematics to students in grades 3 to 6. It combines cooperative learning with individualized instruction. Group members work at their own rates, but they are responsible for checking each other's work and help each other learn the material (Slavin, 1995).

Teams-Games-Tournament (TGT)

TGT is a cooperative learning method that uses the same teacher presentations and group work as in STAD, but replaces the quizzes with weekly tournaments, in which students play academic games with members of other teams to contribute points to their group scores (Slavin, 1995).

Test of English as a Foreign Language (TOEFL)

TOEFL measures the ability of nonnative speakers of English to use and understand English as it is spoken, written, and heard in college and university settings (Educational Testing Service, 2005).

Total Physical Response (TPR)

TPR is a language teaching method based on the coordination of language and action. Learners carry out a series of physical actions while they listen to instructions in the target language (Nunan, 1999).

Zone of Proximal Development (ZPD)

ZPD is the distance between the present level of development as indicated by what a learner can do without assistance and the level of potential development as indicated by what a learner can accomplish with assistance from either peers or adults (Vygotsky, 1978).

Organization of the Study

Chapter One contains the introduction, theoretical framework, purpose of the study, research questions, significance of the study, and definition of terms. Chapter Two presents a review of related literature and research pertaining to the problem being investigated. The methodology and procedures used to collect and analyze data for the

study are presented in Chapter Three. The results of data analyses are included in Chapter Four. The last chapter presents a summary of the study, discussion of findings, implications derived from the findings, delimitations and limitations of the study, as well as recommendations for further research.

CHAPTER TWO

REVIEW OF LITERATURE

The purpose of the study was to investigate the effects of cooperative learning in the context of a college EFL course using cognitive and motivational measures. This chapter presents a review of literature pertaining to cooperative learning in ten areas: theoretical underpinnings of cooperative learning, the essence of cooperative learning, differences between cooperative learning and collaborative learning, brief history of cooperative learning, research on cooperative learning in the United States, research on cooperative learning in second language instruction, need for innovation in Taiwan's EFL education, research on cooperative learning in Taiwan's EFL classrooms, second language acquisition and second language learning, as well as pedagogical practices for grammar teaching.

Theoretical Underpinnings of Cooperative Learning

According to Slavin's (1995) model of cooperative learning, cooperative learning is supported by two major categories of theories: motivational theories and social cognitive theories. The following sections explore these two categories of theories pertaining to cooperative learning.

Motivational Theories

This section discusses the most important contemporary motivational theories, including expectancy-value theories, goal theories, and self-determination theories. Some of them are interrelated with each other while others stand alone.

Expectancy-Value Theories

Expectancy-value theories are a set of conceptualizations contributed by many scholars over the course of half a century. As Wigfield (1994) summed up, the theories conceive that one's motivation to perform a learning task depends on two dimensions: "expectancy of success" in the given task, and the "value" attached to successfully performing the task.

Expectancy of success is related to three factors: (a) how a learner attributes his or her past success or failure; (b) how a learner construes competence; and (c) how a learner maintains self-esteem. These factors are discussed in more details in attribution theory (Weiner, 2000), self-efficacy theory (Bandura, 1993) and self-worth theory (Covington, 1992) respectively. Although the focuses of the theories vary slightly, they are intertwined, and together they illustrate a picture of individuals' performance expectations and their confidence levels in undertaking tasks. As Dornyei (2001) put it, they answer the question of "*Can I do this task?*" (p. 21).

Attribution theory. Attribution theory (Weiner, 2000) assumes that people's motivation is influenced by the "causal stability" of their attributions of past successes and failures. For example, after there is an exam outcome, students might ask themselves what has caused that outcome. The "causal stability" of a given attribution affects their future motivation. "Causal stability" is the duration of a cause. The more "stable" and "constant" it is, the more "uncontrollable" it becomes for learners. On the contrary, the more "unstable" and "temporary" it is, the more "controllable" it becomes for learners. Causes such as luck and level of effort are perceived to be unstable. They are temporary phenomenon and subject to alteration. Therefore, they tend to have less impact on

students' future motivation and behavior. Conversely, causes such as inherent ability and unfairness of a teacher are perceived to be stable and constant. Since they are conceived permanent, the learning situation hence becomes "uncontrollable" to the students. When a cause is seen as constant, a student feels lack of control and lack of power to alter the situation. This type of cause tends to impact motivation negatively and trigger future failure.

Self-efficacy theory. Bandura's (1993) self-efficacy theory is closely related to Weiner's attribution theory. This theory assumes that the way people define competence will influence their interpretation of a learning outcome as well as motivation and future actions on learning. Those who deem competence as an "acquired" skill tend to evaluate their own ability by their personal improvement; those who deem competence as an "inherent" ability tend to evaluate their ability by comparing it against the success of others. Consequently, when encountering frustrations, the former often choose to examine the processes such as effort and use of strategies (i.e., "unstable, "temporary" and "controllable" causes based on the attribution theory) and tend to be ready for more challenges that would broaden their repertoire of knowledge. On the other hand, the latter see their inherent competences (i.e., "stable," "constant" and "uncontrollable" causes based on the attribution theory) as the source of failure; in order to save their self-esteem, they often choose to put forth little effort or select easier tasks so that they could attribute failure to lack of endeavor instead of admitting their low inherent ability. The researcher of the present study has summarized Bandura's (1993, p. 144) remarks on self-efficacy and presents the summarization in a comparison chart (see Table 1) to show the diverse effects self-efficacy could bring upon people who construe efficacy differently.

Table 1
Effects of Self-Efficacy

People Who Construe Efficacy as Inherent	People Who Construe Efficacy as Acquired
➢ Low self-efficacy	➢ High self-efficacy
➢ Difficult tasks = Personal threats	➢ Difficult tasks = Challenges to be mastered
➢ Shy away from difficult tasks	➢ Approach difficult tasks
➢ Have low aspirations and weak commitment to goals	➢ Maintain strong commitment to goals
➢ Maintain a self-diagnostic focus rather than concentrate on how to perform successfully	➢ Maintain a task-diagnostic focus that guides effective performance
➢ Dwell on personal deficiencies, possible obstacles, and all kinds of adverse outcomes in the face of difficulties	➢ Enhance and maintain efforts in the face of difficulties
➢ Slacken their efforts and give up quickly in the face of failure	➢ Attribute failure to insufficient effort or deficient knowledge and skills that are acquirable
➢ Slow to recover sense of efficacy after failure or setbacks	➢ Quickly recover sense of efficacy following failures or setbacks
➢ Fall easy victim to stress and depression	➢ Have low vulnerability to depression

Self-worth theory. Covington's (1992) self-worth theory is also closely related to Weiner's attribution theory. It assumes that the utmost human priority is the quest for self-acceptance and that, in order for people to believe that they have worth as a person in the school context, they need to believe they are academically competent first. Therefore, they often choose to enhance or at least protect their sense of academic competence in order to sustain their sense of self-worth. This is where attribution theory enters the picture. In order to enhance one's sense of control and sense of self-worth, a learner's

most preferred attribution for failure is to evade trying, whereas the most preferred attributions for success are ability and effort (Covington & Omelich, 1979).

Model of Task Values. While the above theories on the "expectancy" dimension answer to the question of "*Can I do this task?*" the second constituent of expectancy-value theories answers to the question of "*Do I want to do the task?*" (Dornyei, 2001). Eccles and her colleagues (Eccles, 1987; Eccles & Wigfield, 1995, 2002; Eccles (Parsons) et al., 1983) have identified four types of task values: attainment value, intrinsic value, utility value, and cost. Attainment value refers to personal importance of performing well on a task. This type of value relates directly to one's ideal self-schemata. For example, if an individual identifies with masculinity and competence in a given area, he will have higher attainment value for a task if it allows him to confirm the self-schemata. Intrinsic or interest value refers to enjoyment and pleasure that an individual gets when performing the task itself. Utility value refers to the degree of relationship an individual perceives between a task and long-term or short-term goals. When people do not particularly enjoy a certain task (i.e., low intrinsic value), they might still do it if they see the task as a mediator to a future goal. Last but not least, cost refers to the negative perspectives an individual attaches to a task, such as fear of the consequence, task anxiety, and the amount of effort required. When people assign high cost to a task, they are more liable to avoid doing the task. The role of educators, therefore, is to design curriculum and instruction that would minimize the fourth type of task value while enhancing the first three types so that learners have sufficient motivation to participate in learning tasks. Table 2 summarizes expectancy-value theories.

Table 2
Summary of Expectancy-Value Theories

Dimension	Answer to the Question of . . .	Component	Focus
Expectancy of success	"Can I do this task?"	Attribution theory	How a learner attributes his or her past success or failure
		Self-efficacy theory	How a learner construes competence
		Self-worth theory	How a learner maintains self-esteem
Value	"Do I want to do this task?	Attainment value	Personal importance of performing well on a task
		Intrinsic/interest value	Enjoyment/pleasure an individual gets when performing a task itself
		Utility value	Degree of relationship an individual perceives between a task and current/future goals
		Cost	Negative perspectives an individual attaches to a task

Goal Setting Theory

Locke and Latham's (1990) goal setting theory claims that human behaviors are regulated by goals or purposes. While a goal, once chosen or accepted by an individual, is not constantly in one's conscious level, it stays in the background and is readily called into consciousness to guide the subsequent behaviors heading for the goal. Based on the theory, difficult goals bring about a higher level of performance than do easy goals; specific difficult goals bring about a higher level of performance than do ambiguous goals of "do your best" or no goal. Moreover, Locke and Latham reviewed several empirical studies and pointed out a number of factors that could impact how one sets goals. Many of these factors are closely related to the theoretical model of cooperative learning proposed by Slavin (1995). The following table presents a summary of Locke

and Latham's literature review on how these factors could influence a learner's goal setting.

Table 3
Factors Influencing Goal Setting and Goal Commitment

Factors	Empirical Findings	Investigators
Role modeling	Observing a high-performing role model brings about higher personal goal setting and higher commitment to difficult goals than observing a low-performing model.	Rakestraw & Weiss (1981); Earley & Kanfer (1985)
Competition	Competition brings about higher personal goal setting (but not higher goal commitment) than no competition on a brainstorming task.	Mueller (1983)
Group goals	Having group goals on top of personal goals brings about higher goal commitment to the personal goals than having personal goals alone.	Matsui, Kakuyama, & Onglatco (1987)
Encouragement	Encouragement and persuasion increase level of goal setting.	Garland & Adkinson (1987)
Feedback	Giving performance feedback brings about higher goal setting than not giving feedback.	Erez (1977)

Self-Determination Theory

As Dornyei (2001) pointed out, one of the most common distinctions in motivational theories is that between extrinsic motivation and intrinsic motivation. Based on the research conducted by Vallerand (1997), over 800 studies dealing with the paradigm have been published. Traditionally, extrinsic motivation is considered a factor that could negatively affect intrinsic motivation (see, for example, Kohn, 1991a, 1991b), but instead of treating them as opposing forces in two polarized categories, self-determination theory (Deci & Ryan, 1985; Deci, Vallerand, Pelletier, & Ryan, 1991) proposes four forms of behaviors that can be placed on an "internalization continuum"

(Deci & Ryan, 1985, p. 137) based on the degree of regulation that is internalized from outside to inside an individual. Due to the basic human psychological needs for competence and self-determination, people are innately motivated

> . . . to master and incorporate many behaviors that are not themselves intrinsically motivated but are valued by the social environment and thus are instrumental for the . . . long-term effectiveness. Behaviors that the organism would not do naturally will have to be extrinsically motivated, but these behaviors may be integrated into the realm of self-determination. (p. 131)

The internalization process transfers people from the right (controlled and extrinsic) toward the left (self-determined and intrinsic) of the continuum in a social context, which, according to Deci and his colleagues (Deci, Vallerand, Pelletier, & Ryan, 1991, p. 329), plays a significant role on the effectiveness of the internalization and integration process. The researcher of the present study has created the following figure to graphically represent the concept.

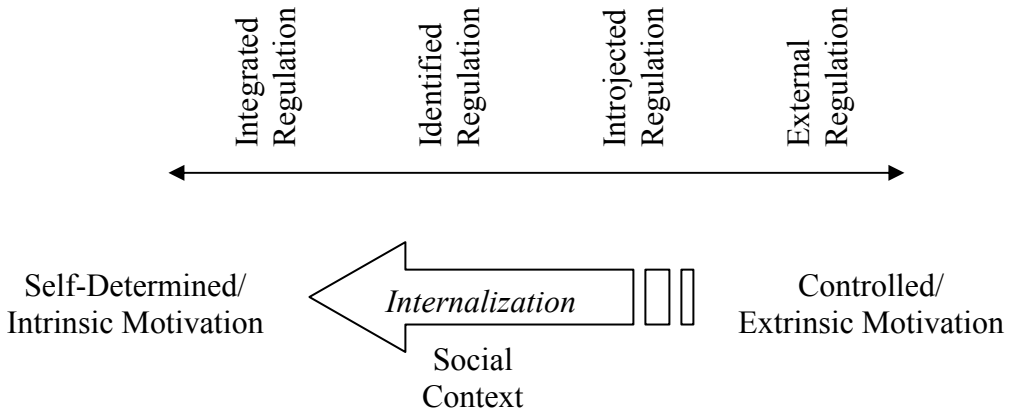

Figure 1. Graphic representation of the self-determination theory.

The four forms of behaviors proposed by Deci and Ryan (1985) are: externally regulated behaviors, behaviors regulated through introjection, behaviors regulated through identifications, and integrated regulated behaviors. Table 4 provides a brief description of these behaviors by the degree of internalization.

Table 4
Motivated Behaviors by the Degree of Internalization

Types of Regulation	Description	Scenario
Externally regulated behaviors	Behaviors that require the presence of external contingencies.	"Mom's mad. I have to clean up the room now."
Behaviors regulated through introjection	Behaviors that occur when individuals monitor their own behaviors but have not yet accepted the regulation as their own. This type of behaviors entails self-control and often involves debates within oneself about to do or not to do.	"I should clean up my room. Good boys clean up their rooms."
Behaviors regulated through identifications	Behaviors that come about when individual have arrived to the state of valuing and identifying the actions.	"Let me clean up my room now—it lets me find things easier."
Integrated regulated behaviors	Behaviors that occur when the value of given actions is already assimilated with one's personal beliefs, needs, and identities.	"I like to clean up my room."

(Deci & Ryan, 1985, p. 135-137)

Each form of the above behaviors contains more self-regulation than the previous form and is one step closer to intrinsically motivated behaviors. It is worth noting, however, that while it appears that Deci and Ryan (1985) initially saw integrated regulation as identical to intrinsic motivation (p. 140), they and their colleagues (Deci,

Vallerand, Pelletier, & Ryan, 1991) later stated that integrated regulation is not yet identical to intrinsic motivation. “Intrinsic motivation is characterized by interest in the activity itself, whereas integrated regulation is characterized by the activity’s being personally important for a valued outcome” (p. 330). Nonetheless, despite the discrepancy, it is clear that self-determination theory emphasizes human motivation as an active internalization process. Intrinsically motivated behaviors originate from oneself and are the prototype of self-determination; extrinsic motivation can facilitate intrinsic motivation due to human beings’ need for self-determination.

Summary on Motivational Theories

This section presents various motivational theories that underpin cooperative learning. Weiner’s (2000) attribution theory assumes that motivation is affected by how people attribute their past success or failure (i.e., stable, constant, and thus uncontrollable factors versus unstable, temporary, and thus controllable factors). Bandura’s (1993) self-efficacy theory maintains that, if individuals deem competence as “acquired” (i.e., controllable based on the attribution theory), they focus on personal improvement and maintain strong commitment to goals. Conversely, when individuals deem competence as “inherent” (i.e., uncontrollable), they maintain a self-diagnostic focus and recoil from challenging tasks in fear of having to acknowledge low inherent ability in case of unsatisfying performance. Similarly, Covington’s (1992) self-worth theory assumes that learners with low confidence often avoid working hard so that they can attribute failure to level of effort exerted (controllable) to retain their sense of control and self-worth.

Based on the motivational theories reviewed, it appears that in order to enhance motivation, instruction needs to be tailored to help learners perceive competence as

acquired skills and to enhance their sense of control over learning tasks. In order for learners to perceive competence as attainable through efforts and to make them believe their power in making a difference, allowing students to make improvement against their own past performance rather than against their classmates seems a reasonable solution. In cooperative learning, this pedagogical practice is called "equal opportunities for success," a feature shared by many cooperative learning methods, including the Student Teams Achievement Divisions (STAD) that was employed in this study. Equal opportunities for success will be discussed in more details in the section of the Essence of Cooperative Learning in this chapter.

Locke and Latham's (1990) goal setting theory argues that human behaviors are regulated by goals and that the setting of personal goals are in turn influenced by factors such as group goals, role modeling, encouragement, and feedback. These factors are compatible with Slavin's (1995) model of cooperative learning (see Figure 2). For example, the goal setting theory argues that having group goals on top of personal goals brings about higher goal commitment to the personal goals than having personal goals alone. Correspondingly, the model of cooperative learning argues that the setting of group goals will trigger motivation to learn, motivation to encourage group members to learn, and motivation to help group members to learn.

While some critics of cooperative learning (e.g., Kohn, 1991a, 1991b) argue that extrinsic motivation triggered by cooperative learning can negatively affect intrinsic motivation, proponents of cooperative learning believe otherwise. Deci and his colleagues' (Deci & Ryan, 1985; Deci, Vallerand, Pelletier, & Ryan, 1991) self-determination theory is apparently very much in line with the perception of

cooperative learning advocates in this regard. The self-determination theory presents four forms of behaviors on a continuum based on the degree of motivation internalization; it clearly argues that extrinsic motivation can facilitate intrinsic motivation and transfer a learner from the right (controlled and extrinsic) toward the left (self-determined and intrinsic) of the continuum.

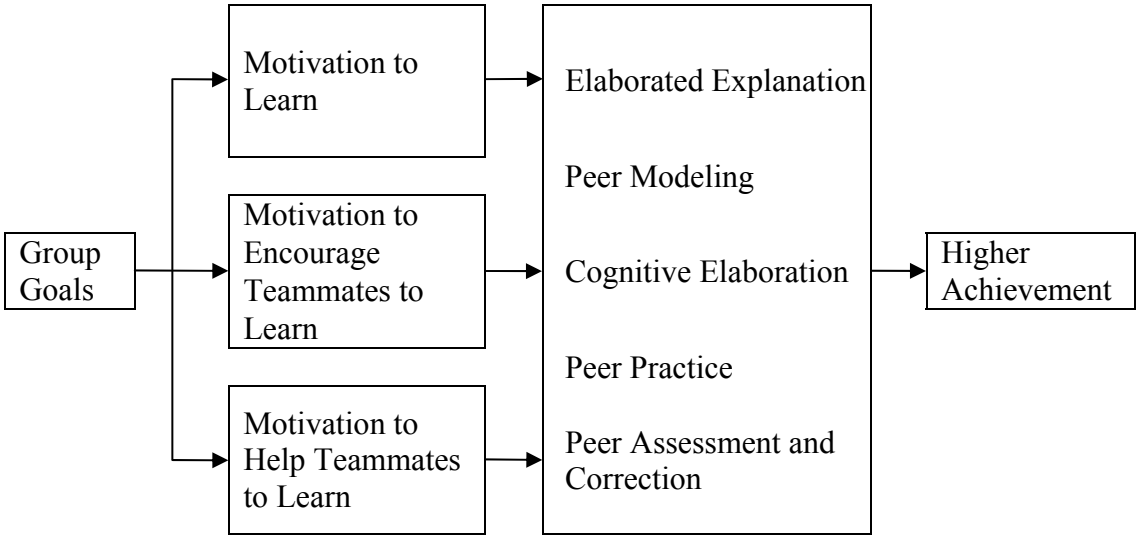

Figure 2. Slavin's model of cooperative learning.

Social Cognitive Theories

Based on Slavin's model (1995), cooperative learning facilitates learning not only because it motivates learners with shared goals but also because it further situates learners in a social context, which provides a stage for cognitive development through elaborated explanations, peer tutoring, peer modeling, cognitive elaboration, peer practice, peer assessment and correction (see Figure 2). As Merriam, Caffarella and Hansman put it, "learning does not occur in a vacuum" (Merriam & Caffarella, 1999, p. 22); rather,

learning is shaped by "the nature of the interactions among learners, the tools they use within these interactions, the activity itself, and the social context in which the activity takes place" (Hansman, 2001, p. 45). It takes the context, culture, and tools in the learning situation for learning to happen. In a similar vein, Perry (1970) pointed out in his scheme of cognitive development that peer interactions help a learner advance from a lower level of cognitive development into a higher level.

Vygotsky

One of the most prominent social cognitive theorists that deem social context significant to cognitive development is Vygotsky. Vygotsky claimed socialization as the foundation of cognition development (1978, p. 57 & p. 90). The internalization of knowledge, according to Vygotsky, is a progression that begins with an interpersonal process before it proceeds into an intrapersonal one; a learner's development first takes place on the social level (between people) before it moves on to the individual level (inside an individual) (p. 57).

Additionally, Vygotsky asserted that socialization facilitates learning because the process of working with others offers a learner an opportunity to operate within his or her "zone of proximal development." Zone of proximal development has been defined as the distance between the current level of development as indicated by what a learner can do *without* assistance and the level of "potential development" as indicated by what a learner can accomplish *with* assistance from either peers or adults (1978, p. 86). The rationale that social interaction with peers enhances learning lies on the fact that collaboration or cooperation with peers lets learners work closely within one another's levels of proximal development. When learners work closely within one another's levels of proximal

development, they can receive explanations that are presented to them in a simpler and more comprehensible fashion than if they were provided by one of a very different mental age.

As a result, the process of cooperation with peers benefits students academically because "what is in the zone of proximal development today will be the actual developmental level tomorrow" (Vygotsky, p. 87). Vygotsky contended that what a learner can accomplish through the tool of social interaction at the moment he or she will be capable of accomplishing independently in the near future. He stressed that social interaction and cooperation with peers are indispensable factors for inner speech and metacognition to take place (p. 90). "[A]ll the higher functions," emphasized Vygotsky, "originate as actual relations between human individuals" (p. 57).

Piaget

Piaget's sociological theory appears to be much less renowned than his individualistic theory. He has been criticized for refuting the significance of the social aspect and thus for having a solely "individualistic" theory of intelligence (Kitchener, 1991). However, a review of literature shows that Piaget does value the significance of social interaction. Kitchener contended that the reason Piaget's sociological theory is not widely recognized by the scholarly world could probably be attributed to the fact that much of it is contained in his untranslated French works.

In discussing moralities and social relations, Piaget (1932) condemned traditional schools, which offer whole-class instruction, competitive examinations, and individual homework. He criticized that the procedure "seems to be contrary to the most obvious requirements of intellectual and moral development" (p. 412). He stated that working in

groups can "correct" the problem and that "cooperation is . . . essential to intellectual progress" (p. 413).

To Piaget, experience is an indispensable element for intellectual development, but he contended that exposing to experience alone is inadequate for learning to take place; the learner has to be "active" in the process (Duckworth, 1964; Piaget, 1964). In a conference on cognitive studies and curriculum development, Piaget (Duckworth, 1964) explained to his audience that a learner could be active either individually or cooperatively:

> When I say "active," I mean it in two senses. One is acting on material things. But the other means doing things in social collaboration, in a group effort. This leads to a critical frame of mind, where [learners] must communicate with each other. This is an essential factor in intellectual development. Cooperation is indeed co-operation. (p. 4)

Piaget (1964) argued that all developments consist of "momentary conflicts and incompatibilities which must be overcome to reach a higher level of equilibrium" (p. 19). Duckworth, a student of Piaget's at the Institute of Genetic Epistemology in Geneva, Switzerland, elaborated on what Piaget means by "equilibration."

> Piaget sees the process of equilibration as a process of balance between assimilation and accommodation in a biological sense. An individual assimilates the world-which comes down to saying he sees it in his own way. But sometimes something presents itself in such a way that he cannot assimilate it into his view of things, so he must change his view-he must accommodate if he wants to incorporate this new item. (Duckworth, 1964, p. 4)

Equilibration involves two complementary activities: assimilation, in which learners use their current schemes to make sense of the external world; and accommodation, in which they modify existing schemes or build new ones after they notice that the existing thinking does not fully capture the reality of the outside world. Equilibration is a process of restoring balance, and this process provides an opportunity for learners to grow and develop (Piaget, 1950). To Piaget (1932), individual activities and group activities both play important roles in the process. "Social life is . . . a complement of individual 'activity'," he argued (p. 413).

In addition to his equilibration theory, Piaget's social transmission theory (1964) provides a rationale for cooperative learning, and it is quite consistent with Vygotsky's zone of proximal development. Piaget argued that learners are receptive to new information only when they are in a state where they are able to comprehend the substance, that is, when they have a structure which enables them to assimilate it. When learners have a structure that enables them to assimilate the information, they are in the zone of proximal development, and working in peer groups often enables learners to help each other move to the next level of development.

Dewey

Dewey (1916, 1963) also deemed participation in social environment as critical to learning. In a similar way that Piaget criticized traditional whole-class instruction, Dewey charged traditional instruction for failing to "secure the active cooperation of the pupil in construction of the purposes involved in his studying" (1963, p. 67). He emphasized that in a cooperative setting, "the individual appropriates the purpose which actuates it, becomes familiar with its methods and subject matters, acquires needed skills, and is

saturated with its emotional spirit" (1916, p. 26). For Dewey, simply waiting passively for the instructor to hand-feed knowledge does not warrant learning; learners need to gain experience through activities in which they actively participate and cooperate with others.

Although Dewey rejected teachers as authoritarian figures, he appeared to be in disagreement with the more extreme advocates of learner-centered progressivism. While Dewey's view of experiential education calls for active (rather than passive) participation of learners, it is worth noticing that Dewey (1964) also stressed the significance of a teacher's active role in the process. According to Dewey, experiential education does not mean that learners get unconstrained freedom in the classroom. The teacher needs to "observe" but not "humor" the interests of students. "To humor the interests is to substitute the transient for the permanent" (p. 179). Instead, the purpose of paying attention to learners' interests is to link them with educative experiences and intellectual development so that essential relations between social experience and human knowledge can be taught and learned effectively.

Dewey (1964) emphasized that it is important for a teacher to keep "constant and careful observation of [learners'] interests" because those interests show "the state of development which the [learner] has reached" (p. 178). Therefore, in experiential education a teacher does not "stand off and look on; the alternative to furnishing ready-made subject matter and listening to the accuracy with which it is reproduced is not quiescence, but participation, sharing, in an activity" (1924, p. 188). According to Dewey (1964), students' interests are always indicative of some power below. A teacher needs to keep "continual and sympathetic" (p. 178) observation of their interests in the process of activities so that he or she can detect what they are geared up for, and what teaching

materials could work most efficiently and productively. In terms of the teacher's role in participating and monitoring the group process, Dewey's view seems to be more aligned with cooperative learning than with collaborative learning (see the discussion on "Differences between Cooperative Learning and Collaborative Learning" later in the chapter).

Bruner

Like Vygotsky, Piaget, and Dewey, Bruner (1990) considered active participation and personal interaction imperative. In his claim about language acquisition, Bruner maintained that early acquisition of a language entails communication. "Language is acquired not in the role of spectator but through use. Being 'exposed' to a flow of language is not nearly so important as using it in the midst of 'doing'" (p. 70). Using and practicing a language in a cooperative group is again a better way of facilitating language learning than listening passively to an instructor because for the complete development of language skills, "they all depend upon being practiced and shaped by use" (p. 72). "It is only *after* some language has been acquired in the formal sense, that one can acquire *further* language as a 'bystander.' Its initial mastery can come only from participation in language as an instrument of communication" (p. 73).

Bandura

Bandura's social cognitive theory (1986) presents an interactive model of causality in which the environment, behaviors, and cognitive and personal factors all function as causal factors of each other (see Figure 3). Bandura termed the relationship as "triadic reciprocality" (p. 23).

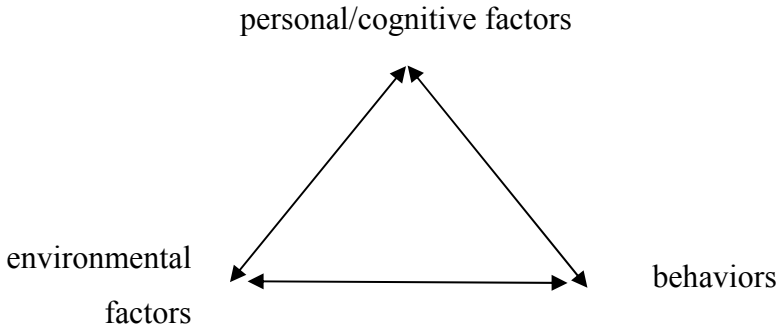

Figure 3. Model of triadic reciprocality.

Based on the model of triadic reciprocality (Bandura, 1986), the causal relationship between the person, the behavior, and the environment is not linear. Instead, the nature of the relationship is interactive. Multiple factors are often required to produce a certain effect. Bandura contended that in the triadic system the three types of factors are "highly interdependent" (p. 24). However, he also noted that the levels of influence exercised by each of the three types of factors vary for different people, different behaviors, and different settings.

According to Bandura (1986), an individual's thoughts and feelings can be shaped, directed, and modified through modeling and social persuasion (p. 25). Modeling not only can direct attention, enhance stimulation, and facilitate learning, but also can arouse emotions. To facilitate development of cognitive skills, verbal modeling of thought processes is necessary. "Cognitive skills are less readily amenable to change . . . when covert thought processes have not been adequately reflected in modeled actions" (p. 74). Bandura explained that if the model solves a problem without demonstrating the thought

process, the observer could see only the end result without the slightest idea how that has been accomplished.

> Learning cognitive skills can be facilitated . . . by having models verbalize their thought strategies aloud as they engage in problem-solving activities. The covert thoughts guiding the actions are thus made observable through overt representation. Modeling both thought and actions has several helpful features that contribute to its effectiveness in producing generalized, lasting improvements in cognitive skills. (p. 74)

It is worth noting, however, that although Bandura argued for the effect of modeling, he also reminded that it takes time for modeling to exert its impact on cognitive skill development. "The production of a reciprocal effect takes time" (p. 25).

Summary on Social Cognitive Theories

Cooperative learning is supported by social cognitive theories proposed by Vygotsky, Piaget, Dewey, Bruner, and Bandura. Vygotsky (1978) argued that socialization is the groundwork of cognition development, and that the process of cooperation or collaboration with peers benefits learners cognitively because it allows learners to work close to one another's zone of proximal development. Piaget's social transmission theory (1964) provides rationale for cooperative learning in a similar way. Piaget contended that individuals are readily amenable to cognitive growth only when they are in a condition where they can understand the concept (i.e., zone of proximal development). Working with peers enables individuals to help each other move to the next cognitive stage. In addition, Piaget's equilibration theory (1932, 1950, 1964) contends that cognitive developments consist of conflicts, which must be overcome

through the process of equilibration, including assimilation and accommodation. Equilibration in turn can be achieved by means of both individual and social activities.

Like Vygotsky and Piaget, Dewey (1916, 1963) and Bruner (1990) considered participation in social environment and interpersonal communication key to cognitive development. Dewey argued that people need to gain experience by actively participating and cooperating with others. Bruner further contended that a language learner needs not only to be exposed to language but also to use and practice the language in a social setting.

Bandura's social cognitive theory (1986) bestows an interactive model of triadic reciprocality in which the environment, behaviors, and cognitive and personal factors all serve as determinants of each other. A learner's thoughts and learning motivation can be directed and shaped through modeling. Bandura's comments on modeling also echoes Slavin's (1995) model of cooperative learning pertaining to the ideas of peer modeling and cognitive elaboration. In addition, Bandura emphasized that it takes time for modeling to exert its impact on cognitive growth. This is one of the reasons the researcher of the current study chose to implement the study for the duration of 12 weeks. Like Bandura, the researcher believes that the effect of modeling and momentum of group dynamics in cooperative learning will not develop fully unless learners have been working with each other over an extended period of time.

The Essence of Cooperative Learning

The following sections present what cooperative learning is, including its characteristics and various methods, the distinction between cooperative learning and collaborative learning, and a brief history of cooperative learning. Subsequent to these,

there will be detailed discussion on the effects of cooperative learning in the United States and in the field of second language acquisition.

Cooperative learning comes in many forms. Among the most widely used and researched cooperative learning methods are student team learning methods, including Student Teams Achievement Divisions (STAD), Teams-Games-Tournament (TGT), Team Assisted Individualization (TAI), and Cooperative Integrated Reading and Composition (CIRC); Jigsaw methods, including Jigsaw and Jigsaw II; and group investigation (G-I) methods, including Learning Together (LT), Co-op Co-op, and Group Investigation (GI). Although there has not been a universal definition of cooperative learning up till now, Olsen and Kagan (1992) have defined cooperative learning as "group learning activity organized so that learning is dependent on the socially structured exchange of information between learners in groups and in which each learner is accountable for his or her own learning and is motivated to increase the learning of others" (p. 8). According to Johnson, Johnson, and Smith (1991), cooperative learning is "the instructional use of small groups so that students work together to maximize their own and each other's learning. . . . To be cooperative, learning groups must be carefully structured" (p. 12).

In addition to the definitions proposed by Olsen and Kagan (1992) and Johnson, Johnson, and Smith (1991), cooperative learning can be further understood through some common features: heterogeneous grouping, positive interdependence, individual accountability, prior training of social skills, group processing, and equal opportunities for success. The first three were characteristics of all the above-mentioned cooperative learning methods; the other three characteristics were shared among certain methods.

Heterogeneous Grouping

The first step of cooperative learning is the formation of heterogeneous learning groups. There can be two types of heterogeneous group formation. The first type is teacher-assigned grouping based on factors such as achievement level and gender. This type of grouping is often adopted by tutoring-oriented cooperative learning methods (also referred to as student team learning methods), including STAD, TGT, TAI and CIRC. The second type is interest grouping, which is often adopted by project-oriented cooperative learning, including Group Investigation and Co-op Co-op.

Positive Interdependence

The next step to ensure the success of cooperative learning is to structure positive interdependence within a cooperative group (Johnson, Johnson, & Smith, 1991; Olsen & Kagan, 1992). Positive interdependence can be established by creating outcome interdependence and process interdependence. Table 5 presents an outline of ways to structure positive interdependence within a cooperative group.

Table 5
Ways to Structure Positive Interdependence within a Cooperative Group

Outcome interdependence	Group goals Group rewards or celebrations
Process interdependence	Role interdependence Resource interdependence

Outcome Interdependence

There are two ways to realize outcome interdependence. One is to establish group goals; the other is to create group rewards or celebrations. Group goals should always be an indispensable part of the lesson. Whatever the content of the goals is, the instructor

makes it explicit to the students that they have two responsibilities: to master the assigned material, and to make sure that all of the group members master the assigned material. Group members "sink or swim together" (Johnson, Johnson, & Smith, 1991, p. 16). One cannot succeed without others being successful. In addition to shared goals, the instructor might give students a group grade for the group's overall performance and bonus points when all the group members live up to a certain standard. Periodical celebrations of the group's endeavor and achievement also enhance the quality of teamwork.

Veenman, Denessen, Van Den Akker, and Van Der Rijt (2005) argued that the group goals and incentive structure of cooperative learning can create circumstances in which students want to help one another because the only way for them to achieve their personal goals is to help their group members succeed. This standpoint has been supported by empirical studies (e.g., Johnson, Johnson, & Smith, 1991; Matsui, Kakuyama, & Onglatco, 1987) which have indicated that the combination of group goals and group rewards/celebrations enhances achievement over group goals alone.

Process Interdependence

There are two types of process interdependence: role interdependence and resource interdependence. To establish role interdependence, an instructor can assign group members complementary roles, such as recorder, checker, encourager, elaborator, taskmaster, and quiet captain (Johnson, Johnson, & Smith, 1991; Olsen & Kagan, 1992). Responsibilities and possible gambits of these complementary roles are presented in Table 6.

Table 6
Possible Role Assignments, Responsibilities, and Gambits

Role	Responsibility	Gambit
Taskmaster	Keeps the group on task; leads the group discussion; makes sure every member contributes and no one dominates the floor	"Have we finished the second paragraph?" "I think the task is . . ." "What do you think, Jeff?" "John, do you agree?" "I would like to hear from Pat."
Checker	Makes sure that everyone knows what is needed to complete the assignment; checks regularly for members' comprehension by asking them to explain what is being learned; makes sure every member understand the material or agree on a decision before proceeding to the next procedure	"Renee, could you please summarize what we've learned so far for us?" "Do we all agree on that?" "Everyone together on this?"
Encourager	Makes sure that the involvement of the whole team and input from each member are valued	"Let's do a team handshake." "Let's all give Richard a pat on the back."
Elaborator	Elaborates information and knowledge to help members learn	"My understanding of the passage is . . ."
Recorder	Records key points of team discussion	"Let me make sure that I record that right."
Quiet Captain	Makes sure the group is not disturbing other groups	"Let's use our 12-inch voices."

(Adapted from Olsen & Kagan, 1992)

To create resource interdependence, limited materials are provided (e.g., one copy per two members, or each member getting part of the required materials) so that group members have to share and work together.

It should be noted, however, that in order to produce higher achievement, resource interdependence should be used only if outcome interdependence is also present. A study conducted by Johnson, Johnson, and Stanne (1989) examined the effects of the combination of goal and resource interdependence, goal interdependence only, resource interdependence only, and neither of them. The results indicated that, among the four

treatment conditions in small groups, the combination of goal and resource interdependence promoted the highest student achievement, while the use of resource interdependence without goal interdependence produced the lowest student achievement. Classroom teachers, therefore, need to be cautious on the use of resource interdependence because resource interdependence does not enhance learning without the existence of group goals.

Individual Accountability

The third essential element for all cooperative learning methods is individual accountability. Individual accountability is present only when each group member is held responsible by other members for putting in a reasonable share to the group's final outcome. Two scenarios could happen if individual accountability is not well-structured. Students could either fail to notice group members' needs for encouragement and support (Johnson, Johnson, & Smith, 1991) or choose to seek a free ride on others' efforts by leaving the task to their group members (see also Kerr & Bruun, 1983). On one hand, this could diminish students' learning motivation; on the other hand, those members who are stuck doing all the work might actually benefit tremendously on the process of taking over the responsibilities at the expense of the free riders (Johnson & Johnson, 1990).

In Slavin's meta-analysis of research on cooperative learning (1995, 1996), individual accountability was found to be pivotal to the success on cooperative learning performance. The simultaneous use of individual accountability and group goals substantially enhanced the effect of cooperative learning. (See Table 10 Breakdown of Effect Sizes by Characteristics of Cooperative Methods and its discussions for more detail.)

Johnson, Johnson, and Smith (1991) suggested using the following methods to structure individual accountability: (1) Keep the group size small. The smaller the group size, the greater individual accountability could be. (2) Give each student an individual test. (3) Randomly call on a student to orally present the group's work in front of the whole group or the whole class. (4) Observe group process and record the frequency of each student's participation. (5) Have the checker in each group check his or her members' comprehension by asking them to explain what has been learned or to elaborate the logic underlying the group's answer. (6) Have students teach what they have learned to their group members.

Prior Training of Group Skills

To achieve group goals, group members need to trust one another, communicate clearly and accurately, avoid misunderstanding, accept and assist one another, and resolve disagreements constructively (Johnson & Johnson, 2003). In order to achieve all these, group skills are indispensable. According to Olsen and Kagan (1992), group skills include acknowledging group members' contributions, valuing group members' contributions, asking group members to provide input, praising group members, checking for agreement, keeping the group on task, keeping conversation quiet, and reconciling discrepancies.

Johnson and Johnson (1990) recommended a few steps for teaching students group skills. First, the instructor is to provide the rationale for using group skills. This may include improvement of group dynamics and extra points for the use of group skills. Then, the instructor is to model how and when to use group skills and ask students to role-play the skills with their group members. Next, students are constantly reminded to

use the social skills they have learned so that they can go through the phases of unnatural enactment and internalize the skills. To expedite the process and maximize the effect, combination of group skill practices and role assignments (see Table 6) is suggested by Olsen and Kagan (1992).

Prior training of group skills is emphasized in Learning Together, Co-op Co-op, and Group Investigation but is not emphasized in student team learning methods (STAD, TGT, TAI, and CIRC) or Jigsaw methods.

Group Processing

Johnson, Johnson, and Smith (1991) defined group processing as "reflecting on a group session to describe what actions of the members were helpful and unhelpful and to decide what actions to continue or change" (p. 22). The purpose of group processing is "to clarify and improve the effectiveness of the members in contributing to the collaborative efforts to achieve the group's goals" (p. 22). Like prior training of social skills, group processing is emphasized in group investigation methods (Learning Together, Co-op Co-op, and Group Investigation) but is not emphasized in student team learning methods (STAD, TGT, TAI, and CIRC) or the Jigsaw methods (Jigsaw and Jigsaw II). Empirical studies (Johnson, Johnson, Stanne, & Garibaldi, 1990; Yager, Johnson, Johnson, & Snider, 1986) found that students in the cooperation with group processing condition had higher academic achievement than students in the cooperation without group processing condition.

Equal Opportunities for Success

To enhance learning motivation, some cooperative learning methods stress equal opportunities for success. In STAD and Jigsaw II, the improvement score system allows

students of all achievement levels to make improvement against their own past performance rather than against their classmates of higher ability levels. In TGT, the tournament system of competing against others of similar past performance gives students of all levels opportunities to contribute maximally to their group scores as long as they try their best. In TAI, the individualized instruction tries to tailor to both high achievers' and low achievers' needs for success. In CIRC, equal opportunities for success are realized when students can make practice in their subgroups and receive feedback on their performance.

Summary on the Essence of Cooperative Learning

The key emphases of the most widely researched cooperative learning methods are summarized in Table 7. These methods are listed under two major categories of cooperative learning as classified by Sharan (1980): the Peer Tutoring methods and the Group Investigation (G-I) methods.

It should be clear to the reader by now that cooperative learning is not putting students at the same table and allowing them to chat occasionally while they perform their individual tasks. Cooperative learning is not assigning a project to a group in which one or few students do all the work while the others do nothing but earn the grade. Nor is cooperative learning assigning a report to a group in which members divide the labor and then each works individually on his or her share only. Cooperative learning has a distinct characteristic of being "carefully structured." For group learning to be truly cooperative, the activity has to be structured in a way that certain cooperative elements not only exist but also co-exist.

Table 7
Key Emphases of Major Cooperative Learning Methods

	Peer Tutoring Methods						G-I Methods		
	Student Team Learning Methods								
EMPHASIS	STAD	TGT	TAI	CIRC	Jigsaw	Jigsaw II	LT	Co-op	GI
Heterogeneous Grouping	●	●	●	●	●	●	●	●	●
Positive Interdependence	●	●	●	●	● [a]	●	●	●	●
Individual Accountability	●	●	●	●	●	●	●	●	●
Prior Training of Group Skills							●	●	●
Group Processing							●	●	●
Equal Opportunities for Success	● [b]	● [c]	● [d]	● [e]		● [f]			

Note. STAD = Student Teams Achievement Divisions; TGT = Teams-Games-Tournament; TAI = Team Assisted Individualization; CIRC = Cooperative Integrated Reading and Composition; LT = Learning Together; GI = Group Investigation. [a] Only process interdependence; no outcome interdependence. [b] Improvement points. [c] Tournament system. [d] Individualized. [e] By subgroup. [f] Improvement points.

Differences between Cooperative Learning and Collaborative Learning

The terms "cooperative learning" and "collaborative learning" have been used interchangeably not only by the general population but also by many educational practitioners, including those who utilize peer learning in their classrooms (see, for example, Adams, 2000; Walling, 1987). The confusion is understandable; while some dictionaries illuminate the different natures of cooperation and collaboration (e.g., Chiu, 2000), many do not. For example, one dictionary defines "cooperate" and "collaborate" as "to act or work together for a particular purpose" and "to work with someone for a special purpose," respectively (Cambridge Advanced Learner's Dictionary, 2003). The two definitions are virtually the same.

Nevertheless, while both cooperative learning and collaborative learning are types of peer learning, they are different in many ways—from their underlying assumptions to emphases to implementations (Bruffee, 1999). Table 8 presents a summary of the differences between collaborative learning and two main types of cooperative learning, i.e., Peer Tutoring and Group Investigation methods.

Whereas cooperative learning originated with the assumption that competition could obstruct learning, collaborative learning originated with the assumption that the hierarchical authority structure could obstruct learning. Due to the different assumptions, cooperative learning emphasizes interdependence and individual accountability to ensure that students work together instead of competing with one another. On the other hand, collaborative learning emphasizes student autonomy over structure.

As a result, cooperative learning involves elements that play either little or no role in collaborative learning. For example, goal and outcome interdependence and students' responsibility for one another are essential ingredients in cooperative learning to ensure that every student is making contributions; but these elements are not stressed in collaborative learning because the autonomous nature of this approach has given students power to decide among themselves how things should be done. Meanwhile, whereas the teacher keeps close observation and intervention to make certain interdependence and accountability take place, these are the least of the collaborative teacher's concern because he or she does not want to jeopardize student self-governance. Another difference between cooperative and collaborative learning is the formation of groups. In cooperative learning, it is systematic and often requires the teacher's preparation

Table 8
Differences between Cooperative Learning and Collaborative Learning

	Cooperative Learning		Collaborative Learning
	Peer Tutoring Methods	Group Investigation Methods	
Underlying assumption	Competition can obstruct learning	Competition can obstruct learning	The hierarchical authority structure can obstruct learning
Emphasis	Interdependence and accountability	Interdependence and accountability	Student autonomy
Group formation	Heterogeneous groups	Heterogeneous groups	Random or interest groups
Positive interdependence	Yes	Yes	No
Individual accountability	Yes	Yes	No
Responsibility	For self and each other	For self and each other	For self
Task and group processing	Only task emphasized	Both emphasized	Only task emphasized
Group skills	Not usually taught	Directly taught	Not usually taught
Procedure	Prescribed/specified by the teacher	Prescribed/specified by the teacher	Students choose task roles & decide among themselves how things should be done
Teacher observation and intervention	Often	Often	Seldom
Assignment	Group-based	Group-based	Individual or group-based

(Adams, 2000; Bruffee, 1999; Johnson, Johnson, & Smith, 1991; Sharan, 1980)

beforehand; in collaborative learning, it is spontaneous and often based on the students' interests or physical proximity in the classroom.

Brief History of Cooperative Learning

The idea of cooperative learning goes far back in history. According to Johnson, Johnson, and Smith's research (1991), the concept of peer learning was described as early as the first century by Marcus Fabius Quintilian, who advocated that peer learning could benefit the students. Quintilian was Rome's leading teacher from about 68 AD to 88 AD (Pappas, 2003).

The idea of peer learning was also described in the Talmud, which explicitly stated the importance of having a learning partner to facilitate learning. The Talmud was the body of Jewish law concerned with both religious and non-religious life (Chiu, 2000). There are two Talmuds, written by two different groups of Jewish scholars: the Babylonian Talmud (c. 600 AD) and the Palestinian Talmud (c. 400 AD) (Pappas, 2003). Although it is not clear which Talmud that Johnson, Johnson, and Smith referred to, the value of peer learning was once again stressed between 400 AD and 600 AD. In the 17th century, Czech educational reformer and religious leader John Amos Comenius (1592-1670) emphasized in his writings political unity, religious reconciliation and educational cooperation (Diggins, 1997; Pappas, 2003). Comenius argued that students would learn by teaching and being taught by other students (Johnson, Johnson, & Smith, 1991).

In the late 18th century, Joseph Lancaster and Andrew Bell opened schools in England that used peer learning groups extensively (Johnson, Johnson, & Smith, 1991). The development of these schools appeared to have marked a milestone for peer learning

because, not long afterwards, the idea of peer learning was brought across the Atlantic Ocean when a Lancastrian school was established in New York City in 1806. Peer learning was emphasized in the early 19th century in the United States during the Common School Movement.

In the last three decades of the 19th century, Colonel Francis Parker, the superintendent of the public schools in Quincy, Massachusetts (1875-1880), strongly advocated the use of peer learning groups in class. Being both enthusiastic and powerful, he was able to attract an average of more than 3,000 visitors yearly to observe his implementation of peer learning (Campbell, 1965, cited in Johnson, Johnson, & Smith, 1991). Subsequent to Parker's efforts, John Dewey (1963) advocated the employment of peer learning in his renowned project method. The methods of peer learning ruled the American education through the turn of the century (Johnson, Johnson, & Smith, 1991).

Approximately the same time Parker was promoting with enthusiasm the use of peer learning, Turner in England and Triplett (1897) in the United States began to compare the effects among competitive, individualistic, and peer learning. Their efforts were followed by investigations in the early 20th century by Mayer in Germany and Ringelmann in France (Johnson, Johnson, & Smith, 1998). Two major studies on peer learning and competitive learning were published in the 1920s and 1930s. Maller's research (1929) was probably one of the earliest laboratory studies on cooperation/collaboration; May and Doob (cited in Deutsch, 1949) reviewed literature on peer learning and competition up until 1937.

Peer learning has had its ups and downs in the American education. After it enjoyed success in the late 19th century and the early 20th century, interpersonal

competition gained ground in American public schools and colleges in the late 1930s (Pepitone, 1980), and interest in peer learning died out. A few decades later, however, when public schools were forced to integrate in the 1960s, the interest in peer learning was rekindled. Peer learning was invited back to the classrooms because educators were seeking ways to construct social integration between minority and majority students and to help improve minority students' academic performance (Olsen & Kagan, 1992).

In the 1970s, several research groups in the United States began independently to develop and examine cooperative learning methods in classroom settings (Slavin, 1991b). These groups included Elliot Aronson and his associates (University of Texas at Austin) who developed the Jigsaw method, David Johnson and Roger Johnson (Cooperative Learning Center at the University of Minnesota) who developed Learning Together, as well as David DeVries, Keith Edwards and Robert Slavin (Center for Social Organization of School at the Johns Hopkins University) who developed Teams-Games-Tournament and Student Teams Achievement Divisions. It was during approximately the same period when another group of researchers in Israel, Shlomo Sharan, Yael Sharan, and Rachel Hertz-Lazarowitz (Tel-Aviv University), refined John Dewey's cooperative model and developed Group Investigation.

It is worth noting that, before 1970, almost all the reported studies on cooperative learning had been college-based. Beginning in the earlier 1970s, nonetheless, the positive effects of cooperative learning attracted K-12 educators' attention. The tide turned. Studies at elementary and secondary levels became robust while those at college level became limited. It was not until after the 1990s cooperative learning at college level

began to regain attention from researchers and educators (Johnson, Johnson, & Smith, 1998).

Research on Cooperative Learning in the United States

A large body of research has compared the effects of cooperative learning and whole-class—competitive or individualistic—instruction. Two of the most comprehensive and substantive meta-analyses on the cooperative learning literature were conducted by Johnson and Johnson (1989) and Slavin (1995). Both meta-analyses have shown that in general cooperative learning produces higher achievement than whole-class instruction.

In order to enhance the credibility of the meta-analyses studies, both synthesis studies established methodological criteria for review. Furthermore, both meta-analyses used vote counting and effect size as measurement to evaluate the effect of cooperative learning on student achievement. Vote counting (Hedges & Olkin, 1985; Jackson, 1980) estimates the size of an overall treatment effect by calculating the proportion of studies showing significantly positive, no difference, and significantly negative outcomes. A limitation of the vote counting method is that it only reflects the direction, not the magnitude, of an effect (Hedges & Olkin, 1985). Johnson and Johnson (1989) and Slavin (1995) offset the limitation by also using Cohen's *d* for effect size statistics.

An effect size is a standard measure of the mean difference between the experimental (cooperative learning in this case) and the control groups (Gravetter & Wallnau, 2004, p. 270). To be more specific, it describes how well the average student in the cooperative learning group has performed compared to the average student in the control group (Gall, Borg, & Gall, 1996). According to the criteria on Cohen's *d*

(Cohen,1988), a study with an effect size of .20 or larger is considered to have medium effect; an effect size of .80 or larger is considered to have large effect. According to Gall, Borg, and Gall (1996, p. 6), an effect size of .33 or larger is usually considered to have practical significance.

Johnson and Johnson (1989) reviewed 539 studies spanning across 93 years (1897 to 1989) with 68% of which conducted within the most recent 29 years. These studies covered a wide array of subject areas; 85% of the studies randomly assigned individuals or groups to treatment conditions; 98% were conducted in North America; 33% were conducted in elementary schools, 21% in secondary schools, 40% in colleges, and 5% on adults.

Johnson and Johnson (1989) categorized the studies into high-, medium-, and low-quality with a systematic point scale, and they measured effect sizes according to the quality of studies. Criteria of categorization included randomization, clarity of the control condition, rotations of experimenters, same curriculum for the experimental and control groups, and verification of the implementation of the independent variable. Results showed that differences among the high-, medium-, and low-quality studies for the cooperative versus competitive comparisons were not statistically significant. The cooperative groups outperformed the competitive groups in all three quality levels with mean effect sizes ranging from .51 to .88. The overall vote counting indicated positive treatment effect of cooperation, with 60% of the studies with significantly positive outcomes, 32% with no difference, and only 8% with significantly negative outcomes. The overall mean effect size was .73. Considering the scope of the studies reviewed, it appears that the positive results of the meta-analysis have substantial generalizability.

The other landmark research synthesis was conducted by Slavin (1995). This review also indicated positive effect of cooperative learning. The inclusion criteria for the meta-analysis were also explicitly stated. The cooperative learning and control groups must study the same material; the cooperative learning and control groups must be equivalent to begin with; study must at least last 4 weeks (20 hours); and achievement tests must measure objectives taught in both groups. Those that did not meet the criteria were excluded. All together there were 90 primary studies qualified for analysis. These studies spanned over the course of 24 years (1972 to 1995).

The meta-analyses first categorized qualified studies into nine cooperative learning methods; then for each method vote counting and mean effect size were calculated to show the effect's direction and magnitude. Table 9 presents the effect sizes by cooperative methods. Some studies compared multiple cooperative learning methods to control groups and were listed more than once. The table therefore presents 99 comparisons of cooperative learning and control methods. Because not all primary studies had effect sizes available, the number of studies counted for mean effect size was slightly less than that of the total studies.

Slavin's research review (1995) further explored the factors that could have affected student achievement by comparing the results of vote counting and mean effect sizes of various elements of cooperative learning. A breakdown of effect sizes by characteristics of cooperative methods is presented in Table 10.

Table 9
Breakdown of Effect Sizes by Cooperative Methods

	Mean ES		Percentage of Studies						
			Significantly Positive		No Difference		Significantly Negative		Total Studies
Student Team Learning:									
STAD	.32	(26)	69	(20)	31	(9)	0	(0)	29
TGT	.35	(7)	75	(9)	25	(3)	0	(0)	12
TAI	.15	(6)	100	(6)	0	(0)	0	(0)	6
CIRC	.29	(8)	100	(8)	0	(0)	0	(0)	8
All STL	**.32**	**(47)**	**77**	**(43)**	**23**	**(12)**	**0**	**(0)**	**55**
Jigsaw	.12	(8)	31	(4)	46	(6)	23	(3)	13
Learning Together	.04	(8)	42	(5)	42	(5)	17	(2)	12
Group Investigation	.06	(6)	50	(3)	50	(3)	0	(0)	6
Structured Dyads	.84	(4)	100	(6)	0	(0)	0	(0)	6
Other	.10	(4)	29	(2)	71	(5)	0	(0)	7
All Studies	**.26**	**(77)**	**64**	**(63)**	**31**	**(31)**	**5**	**(5)**	**99**

Note. Numbers in parentheses are total numbers of studies in each category.
(Adapted from Slavin, 1995)

Table 10
Breakdown of Effect Sizes by Characteristics of Cooperative Methods

	Mean ES		Percentage of Studies						
			Significantly Positive		No Difference		Significantly Negative		Total Studies
Group Goals and Individual Accountability	.32	(52)	78	(50)	22	(14)	0	(0)	64
Group Goals Only	.07	(9)	22	(2)	56	(5)	22	(2)	9
Individual Accountability	.07	(12)	35	(6)	47	(8)	18	(3)	17
No Group Goals or Individual Accountability	.16	(4)	56	(5)	44	(4)	0	(0)	9
All Studies	**.26**	**(77)**	**64**	**(63)**	**31**	**(31)**	**5**	**(5)**	**99**

Note. Numbers in parentheses are total numbers of studies in each category.
(Adapted from Slavin, 1995)

As shown in the table, cooperative methods that integrated group goals and individual accountability generated substantially higher positive effect than the methods that employed group goals only, individual accountability only, or neither. Notice that 78% of studies that employed both group goals and individual accountability found significantly positive effects, and there was no significantly negative effect. On the other hand, only 22% of studies that employed group goals alone found significantly positive effects, and an equivalent percentage of studies found significantly negative effect.

In addition to the results that vote counting revealed, mean effect size also shed light on the importance of the simultaneous use of group goals and individual accountability. As presented in the table, the mean effect size across the 52 studies that incorporated both group goals and individual accountability was .32. But the mean effect size across the 9 studies that employed group goals alone was only .07. (See the previous Individual Accountability section for more discussion on lacking of individual accountability.)

The studies that were reviewed in Slavin's meta-analysis (1995) were mostly conducted in the United States, with a few exceptions conducted in Israel, the Netherlands, and Nigeria. Grade levels in the studies ranged from K through 12. Although the magnitude of the findings in Slavin's review seems less substantial than that of Johnson and Johnson's review (1989), based on Cohen's (1988) and Gall, Borg, and Gall's (1996) criteria on effect size, the findings still present cooperative learning as an effective approach for the academic achievement of students.

Research on Cooperative Learning in Second Language Instruction

According to Olsen and Kagan (1992), cooperative learning provides second language (L2) students more opportunity for language development than traditional language classes do. They argued that, quantitatively, cooperative learning amplifies active use of language when L2 students try to comprehend or produce the language within their cooperative groups; qualitatively, cooperative learning increases linguistic complexity as L2 learners try to reiterate, explain, expand, and elaborate their thoughts to request clarification or to elucidate their points. Nevertheless, although the effectiveness of cooperative learning has been extensively studied in a wide array of subject areas for many decades, the concept of cooperative learning was not introduced to the arena of L2 instruction until 25 years ago (Gunderson & Johnson, 1980). Even after Gunderson and Johnson (1980), cooperative learning did not spark much interest in the L2 field for another 10 years or so.

In the late 1980s, Bejarano (1987) examined the effects of cooperative learning methods, including Student Teams Achievement Divisions (STAD) and Discussion Group, versus whole-class instruction. This study has since been quoted by a good number of researchers (e.g., Chang & Smith, 1991; Ghaith, 2003a; Liang, 2002; Olsen & Kagan, 1992) interested in the effect of cooperative learning in the L2 field and is in a sense considered a landmark study. Nevertheless, the researcher of the present study argued that the value of Bejarano's (1987) study lies on its being one of the earlier studies that examined the effect of cooperative learning in L2 classrooms rather than its empirical implications. The empirical value of the study has been seriously discounted

due to its methodological weaknesses in three areas: testing instrument, pretest, and data analysis.

In Bejarano's (1987) study, no validity and reliability information on the achievement test was reported. Although a pretest was given, some students did not take it. In the STAD groups, the participating teachers used "whatever information is available (e.g., test scores, grades, or personal judgment)" (p. 486) to rank and group the students. According to the model of STAD (see Slavin, 1995), grouping should be made based on the principle that each group on average has equal ability level compared to any other group in class. (One of the reasons is that they will be compared with other teams for group recognition.) Evidently the teachers were unable to use the pretest to group the students since some of the students missed taking it. If a teacher chose to use "personal judgment" to rank the student and to assign students the "base score," the study might very likely be contaminated.

Personal judgment could very likely be subjective. A low achiever might get a base score that was too high for his true level of ability and suffer from having difficulty to earn "improvement score." Or a student might get a base score that was way too low to his true ability level, e.g., an average student was mistakenly perceived as a lower achiever by the teacher and was given a base score of 40 instead of 65. He might find it too easy to earn "improvement score" and therefore lacked incentives to seek his teammates' help or work really hard. In the meanwhile, his teammates might think his "improvement scores" were high enough and ignore the fact that he was actually not making progress at all if his true starting ability level was taken into consideration. In any of these cases, it could result in complaint of unfairness or lack of motivation, and

therefore negatively affected the reliability of the study results. Even in a situation where personal judgment was the only resolution, criteria of evaluation needed to be set up very clearly so that the judgment could be as objective as possible.

Probably because of incomplete pretest data, Bejarano's (1987) study was unable to use the pretest data as covariance and thus had no control of potential pretreatment differences between the cooperative learning and whole-class groups. In addition, instead of using three analyses of variance (ANOVAs), post hoc should have been used to discover where difference among the pairs of group means lies and to maintain an overall alpha level.

In the 1990s, research on cooperative learning in second language instruction began to gain some momentum. Three edited books (Holt, 1993; Kessler, 1992; Lantolf & Appel, 1994) with a good collection of theoretical and empirical studies were published along with a number of journal articles (e.g., Calderon, Hertz-Lazarowitz, & Slavin, 1998; Milleret, 1992; Szostek, 1994). One of the most robust research studies among them is a two-year study conducted by Calderon, Hertz-Lazarowitz, and Slavin (1998). The study scrutinized the effects of a cooperative learning program, Bilingual Cooperative Integrated Reading and Composition (BCIRC), on Spanish and English reading, writing, and language achievement. The study stood out among others because of the rigor of its research design, extended length of the study, and meticulousness in its data analysis. The study indicated that, while students in the BCIRC and traditional groups performed at the same level on second grade Spanish reading and third grade English language, those in BCIRC performed significantly better in second grade Spanish writing and third grade English reading. The study also indicated that the longer students

had been in the BCIRC program (i.e., 2 years versus 1 year), the bigger the effect size was over the traditional program (i.e., a large effect size of +.87 versus a medium effect size of +.33).

In the recent years, Ghaith and his colleagues have made a series of efforts to examine the effects of various cooperative learning methods on EFL students (Ghaith, 2001, 2002, 2003a, 2003b, 2004; Ghaith & Bouzeineddine, 2003; Ghaith & El-Malak, 2004; Ghaith & Yaghi, 1998). Their studies, like that of Calderon, Hertz-Lazarowitz, and Slavin's (1998), also demonstrated rigor of scientific inquiries. Attention in detail was given from research design to data collection and analysis. Fidelity of treatment was carefully ensured (e.g., Ghaith, 2003a). Cooperative learning methods scrutinized in these studies included Jigsaw II, Learning Together, and Student Teams Achievement Divisions (STAD). Data collection methods employed included paper-and-pencil tests as well as questionnaires.

Results of the studies indicated that EFL students in cooperative learning groups performed either significantly better or at the same level compared to those in whole-class groups. One study showed that the EFL high school students in Learning Together demonstrated higher academic gains than those in the whole-class instruction (Ghaith, 2003a). Another study showed that, while the EFL college students receiving Jigsaw II performed at the same level on literal reading comprehension as their peers receiving whole-class instruction, the Jigsaw II group significantly outperformed the whole-class group in higher-order reading comprehension (Ghaith & El-Malak, 2004).

In addition to the results favoring cooperative learning in the cognitive domain, the studies indicated that students receiving cooperative learning also appeared to feel

more academic and personal support from their peers and teachers, less school alienation (Ghaith, 2002), and more class cohesion and fairness of grading (Ghaith, 2003b). Between higher achievers and lower achievers, the latter especially enjoyed the personal and academic support they received from their cooperative learning experience (Ghaith & Bouzeineddine, 2003). Likewise, in a recent study conducted by Sellers (2005), the learners showed strong sense of group, reduced anxiety, and enhanced motivation after receiving second language instruction through cooperative learning. The cooperative learning method employed in the study was Co-op Co-op. Data were collected via individual interviews, focus group interviews, questionnaires, students' reflection papers, and course evaluations.

The review of literature in the previous section has shown the benefit of cooperative learning across a wide array of subject areas and age groups. However, the large body of research mainly focused on L1 learners. This section has focused the scope of literature review on how cooperative learning has worked for L2 learners. It appears that cooperative learning could be beneficial to L2 learners in cognitive, social-affective, and linguistic domains.

Need for Innovation in Taiwan's EFL Education

While English has become a global language and while the Taiwanese government puts emphasis on English education, many students in Taiwan are really challenged to learn this foreign language. This section aims to take the reader into Taiwanese EFL classrooms, to understand what has possibly gone wrong, and to recognize the need to seek a possible solution.

While students in the Western countries are often encouraged to play an active role in class by asking and answering questions, the traditional educational framework in Taiwan is not designed to sustain active communicative process. In Taiwan, EFL instruction is still mostly whole-class, teacher-centered rote grammar-translation (Babcock, 1993; Su, 2003; Yu, 1993), and of large class enrollments (Babcock, 1993). Yu described typical English classes in Taiwan as follows,

> If you have a chance to visit EFL classes in Taiwan, you will find that students sit in straight rows and are assigned seats with little or no opportunity for interaction. Students and teachers alike are reinforced for quiet classrooms, despite the fact that an atmosphere constricted with silence is a deterrent, not an aid, to learning. Students tend to remain silent as much as they can. (p. 216)

Su's two-year nationwide investigation (2003) reported that more than 80 percent of English instruction in Taiwan has adopted the grammar-translation method. The grammar-translation method places emphasis on the teaching of grammatical rules and sentence structures of English using Chinese translation. The teacher is the center of the classroom giving instruction with little input from students. Su's data were collected through questionnaire surveys, face-to-face interviews, telephone interviews and classroom observations.

Babcock (1993) argued that the traditional educational model not only deprives students of a natural language learning environment but also their opportunities to express themselves in the target language. Babcock observed college EFL classes in Taiwan and reported that, when students were called upon by an instructor to answer questions, what prevailed in the classroom was often "awkward class silence," "downcast

eyes," "high levels of stress," "acute embarrassment," "loss of face," and "a sense of impending failure" (p. 7).

In Cheng's study (1998), many student interviewees reported feelings of stress in a traditional English classroom because they were worried that their peers, whom they considered rivals in competition for good grades as well as the teacher's attention and approval, might make fun of them if they failed to provide the accurate answer in front of the whole class. Cheng noticed that students' concern over loss of face and their sense of need to compete with and surpass their peers have triggered a vicious cycle, which often increased the anxiety level and caused a chain reaction of poorer and lower self-esteem. Additionally, Cheng's study has shown that the traditional method, which has been adopted by a majority of the English teachers in Taiwan, could bore the students and even totally ruin some students' learning motivation.

In a similar vein, Lai (2001), Chen (1998) and Yu (1993) argued that the traditional method of one-way communication fails to motivate Taiwan's EFL students. In Lai's questionnaire study, students reported relative dissatisfaction with the current college English instruction (M = 2.59 in a five-point Likert scale). They found the design of the current college classes non-motivating (M = 2.59). According to Yu's study, 60 to 70 percent of students have lost their interest in learning English in a traditional EFL classroom. Additionally, Shen (2002) reported Taiwanese students' scores on the Test of English as a Foreign Language (TOEFL) have dropped considerably to the bottom of the international list. Yu (1993) argued that the decrease in English proficiency and learning motivation has become "apparent crises" (p. 217) in EFL education in Taiwan. While acknowledging multiple factors that might also undercut students' motivation, Yu

contended that the traditional teaching method that prevails in Taiwan has produced competition that impacts EFL learning negatively, especially for slower learners. Both Yu and Cheng pointed out the need for Taiwan's EFL instruction to change from teacher- to student-centered learning. "We want students to be more actively involved in their learning, to learn in a healthy learning situation, to develop their sense of interdependence. . . . So, cooperative learning is worth trying" (Yu, 1993, p. 219).

The above literature review indicates a need for change in Taiwan's English instruction. It has motivated the researcher of the present study to find out what is available other than the whole-class instruction. A further literature review has brought the researcher to cooperative learning, which has been widely used in the United States across different subject areas and grade levels and in recent years, in some L2 classrooms in the Middle East.

Research on Cooperative Learning in Taiwan's EFL Classrooms

The benefits of cooperative learning found from the literature reviews presented in the earlier sections have made the researcher of the present study curious about the employment of this approach in Taiwan. Unfortunately, a diligent search has yielded limited studies that sought to understand the implementation and effectiveness of cooperative learning. The studies used in this literature review were located via a comprehensive search of the literature. Electronic searches were performed on ERIC, Dissertation Abstracts, EBSCO Education, and PerioPath databases.

Although there was a relatively small body of research on the implementation of cooperative learning in Taiwanese English classes, a review of the existing literature seems to suggest the approach as a possible alternative to the traditional, whole-class English

instruction. Numerous researchers have suggested that cooperative learning would promote better achievement, higher motivation, and improved social relations (see, for example, Chang, 1995; Chu, 1996; Lo, 1998; Wei, 1996; Yu, 1993). Three recent studies have provided empirical support for this contention (Chen, 1998; Chen, 1999; Liang, 2002). Of these, the study by Liang is the strongest, employing multiple methods of data collection and data analysis (i.e. questionnaires, observations, interviews, testing, and content analysis).

Studies that Employed Interviews

Among the studies this meta-analysis reviewed, the studies by Chen (1998) and Liang (2002) conducted interviews. In the first study, students' voices from both high achieving and low achieving groups were heard through interviews conducted by the teacher as researcher. While a teacher-as-researcher research design gives a study a close-up observation, it might risk sample bias and objectivity in ways of participant selection and participants' over-identification with the researcher (see Glesne, 1999). The second study's data collection process was robust. The researcher interviewed both the teacher and students. The interviews were transcribed verbatim and crosschecked with the interviewee for content validity.

The results of the interviews in both studies have shown cooperative learning as an approach that promotes active participation, higher self-esteem, and lower anxiety for both high achievers and low achievers.

Studies that Employed Questionnaires

Among the studies this meta-analysis reviewed, the studies by Chu (1996), Lo (1998) and Wei (1996) used Likert-scaled questionnaires to understand students'

perceptions on cooperative learning. Lo's questionnaires also included open-ended questions. The results of the questionnaires yielded similar conclusions to the interviews mentioned in the previous section. While the data showed clearly that students had relatively positive attitudes to cooperative learning, an extension on the length of time devoted to cooperative learning would have enhanced the validity of the second study, which devoted only 6 hours throughout a semester to cooperative activities. Longer hours devoted to cooperative activities could reduce risk of the Hawthorne effect, which might take place if the participants' perceptions were influenced by the novelty of the approach.

The first and the third studies were robust in this regard. The questionnaires were administered after the students had received the cooperative treatment for a school year or a semester, respectively. Since it had become a routine when they answered the questionnaires, there was no concern for the Hawthorne effect.

Study that Employed Observations

Liang's study (2002) measured students' linguistic competence and discourse competence through content analysis on their oral tests and observations on their non-verbal cues. A set of carefully defined criteria was created. Content validity and inter-rater reliability were established for the tests and observations.

The study has shown that the cooperative group outperformed the traditional group in their oral performance. The cooperative group also demonstrated better discourse competence by using more discourse markers of openings, transitions, and pre-closings in their conversation, as well as more eye contact and few signs of nervousness.

Studies that Employed Paper-and-Pencil Tests as Instrument

The four studies that employed paper-and-pencil tests to measure the effects of cooperative learning (Chang, 1995; Chen, 1999; Liang, 2002; Wang, 2001) bore mixed results. The first and the fourth studies indicated higher performance after the students received cooperative learning treatment. The second study showed that the cooperative group significantly outperformed the whole-class group in the overall English achievement test (medium effect size of .55) and the cloze test (large effect size of .94). The third study reported no significant difference between the two groups on the paper-and-pencil tests. These mixed results could be due to the following reasons:

(1) *Research design.* In Chang's study (1995), two approaches were implemented in one semester to the same group of students as opposing treatments: whole-class teaching and cooperative learning. Then students were instructed with the whole-class approach, and they received a second test. After they were instructed with the cooperative learning approach, they received another test. The results of the two tests were then compared to examine the effect of cooperative learning. The repeated-measures design risked carryover effect and progressive error, making the results inconclusive. More specifically, while the difference between the two tests could be due to different effects of two treatments, it is also likely to have been caused by the lingering aftereffects of the first teaching approach (i.e., carryover effect). On the other hand, the difference could simply be a result of maturation or, in other words, a function of time (i.e., progressive error) (Gravetter & Wallnau, 2004). Whatever the case was, the research design has made it difficult to determine what the true cause was. Wang's study (2001) had similar problems. The experiment was conducted between the midterm examinations of two

consecutive semesters. The repeated-measures design also risked carryover effect and progressive error based on the same rationale mentioned above. On top of that, there was a break between the semesters before the subjects continued receiving the treatment. Extraneous variables could intervene and threaten the internal validity of the study during the break. For example, students might study hard, watch many English videos, or go to English cramming schools and thus improve their language proficiency. On the other hand, the effect of the teaching strategy could be minimized by the break. In either case, the data could be contaminated. The break was usually 4 or 8 weeks long, depending on the academic calendar the college followed.

(2) *Extent of treatment implementation.* In Chen's study (1999), the experimental group received Student Teams Achievement Divisions (STAD) for 2 months. Two months seem short for a STAD study. STAD relies mainly on group work. Group members help each other and are responsible for each other's success or failure. In order to implement positive interdependence of goals, which is the key to the success of STAD according to Slavin's model (1995), group members need to believe that they and their group members "sink or swim together" (Johnson, Johnson, & Smith, 1991, p. 6). After group identity is formed, it then takes time to develop social interaction and trust within a group. It is not an easy task to accomplish all these within 2 months. When Chen's study concluded that the STAD group significantly outperformed the control group after only 2 months' treatment, it presented STAD as a quite promising method. Nevertheless, one could challenge the difference as a result of the Hawthorne effect. Will the method maintain its effectiveness after it is implemented for a prolonged period? Will it produce

the same result once students see it as a routine rather than a novelty? Chen's study is worth replicating to explore the effect of the same approach in a longer term.

(3) *Instrumentation*. Validity and reliability of the tests were not established.

(4) *Statistical procedures*. Multiple t-tests were used (Chen, 1999; Liang, 2002) to make comparisons among tests. The statistical procedures inflated the chance of Type I error. If the analysis of covariance (ANCOVA) could be used in place of the multiple t-tests, the chance of committing Type I error could be reduced and the conclusion more robust.

Summary of the Studies by Data Collection Methods

The above-mentioned studies collected data through interviews, questionnaires, observations, and testing. Results that were obtained via the first three methods suggest cooperative learning as a more favorable approach than the whole-class approach for students' self-esteem, motivation, and achievement. Results that were obtained via paper-and-pencil tests, however, yield inconsistent conclusions. There is a need for further research to employ improved methodological procedures, including research design, extent of treatment implementation, instrumentation, and statistical procedures, to retest the effects of cooperative learning in comparison to the traditional, whole-class approach.

Review of Studies by Types of Cooperative Methods and Subject Characteristics

Cooperative learning methods come in many forms. According to Sharan (1980), cooperative learning methods can be classified into two major categories: Peer Tutoring methods and the Group Investigation (G-I) methods. Of the studies reviewed, four studies (Chen, 1998; Chen, 1999; Chu, 1996; Liang, 2002) employed the Peer Tutoring methods,

and five (Chang, 1995; Liang, 2002; Lo, 1998; Wang, 2001; Wei, 1996) employed the G-I methods.

Of the four studies that employed the Peer Tutoring methods (Chen, 1998; Chen, 1999; Chu, 1996; Liang, 2002), the samples consisted of junior high, high school and college students. Apparently there is a gap in the research literature regarding the implementation of the Peer Tutoring methods on the elementary school level. In addition, while effects on both academic achievement and social-affective variables have been explored on the junior high and high school levels, the effects of the Peer Tutoring methods on academic achievement have yet to be explored on the college level.

Of the five studies (Chang, 1995; Liang, 2002; Lo, 1998; Wang, 2001; Wei, 1996) that employed the G-I methods, the samples consisted of junior high and college students. There is a gap in the literature regarding the implementation of the G-I methods on the elementary and high school levels.

Furthermore, due to the methodological flaws that have been discussed in earlier sections, there is a need for future research to retest the effects of both the G-I model and the Peer Tutoring methods, using improved methodological procedures, on all grade levels.

Second Language Acquisition and Second Language Learning

The following sections discuss selected theories pertaining to second language acquisition and learning. The first section examines second language (L2) theories that focus on the importance of natural input. The second section examines L2 theories that emphasize consciousness-raising. Based on the findings of some empirical research studies, however, it appears that L2 learners need an integrated approach encompassing the merits

of both orientations to maximize their learning. The third section presents a review of these studies and suggests the combinational use of both form-focused and communicative approaches.

Theories Emphasizing Natural Input

Krashen has developed a wide-ranging second language acquisition (SLA) theory over the years (1982, 1985, 1988, 2002, 2003). While his theory on SLA is celebrated by many scholars in the field, probably an equal number of the scholars are in disagreement. Two of Krashen's most renowned hypotheses—the acquisition-learning distinction hypothesis and the input (comprehension) hypothesis—have been criticized for being ambiguous, circular, contradictory, biased, or lacking in evidence. The following section presents the two hypotheses and their criticisms.

Acquisition-Learning Distinction Hypothesis

Krashen (1976, 1985, 1988, 2002, 2003) claimed that there are two distinct and separate processes to develop L2 competence: acquisition and learning. He defined "acquisition" as "a **subconscious** process identical in all important ways to the process children utilize in acquiring their first language" and "learning" as "a **conscious** process that results in 'knowing about' language" (emphases added) (1985, p. 1). In other words, acquisition takes place in a natural communication context where people are not aware that they are "picking up" (2003, p. 1) the language. On the contrary, learning takes place in a context where conscious attention is paid to rules and error detection. Krashen stated that whereas informal environments help language acquisition, formal environments such as classes facilitate both language acquisition and language learning (1976, 1988). Krashen further argued that learning does not turn into acquisition, and that acquisition is vital for

both children and adults approaching a second language (1976, 1982, 2002), while conscious learning is "very limited" (2003, p. 1) and "peripheral" (2002, p. 213) in the development of second language abilities. For children, stated Krashen (1976), "explicit tutelage is unnecessary" (p. 163).

Criticisms. While Krashen's acquisition-learning distinction hypothesis are acknowledged by many in the field of second language, it is also widely criticized. As McLaughlin (1987) pointed out, it is difficult to differentiate between acquisition and learning. Although with his colleagues Krashen seemed to have operationally defined learning as conscious decisions of grammaticality based on "rule" and acquisition as subconscious judgments based on "feel" (Krashen, Butler, Birnbaum, & Robertson, 1978, p. 82), he has failed to present ways that can determine when a particular process engages learning and when it engages acquisition. McLaughlin (1987) argued that because the acquisition-learning distinction is ambiguous, it is empirically unfeasible to judge whether a learner is functioning based on acquisition or learning in a given situation. As a result, a key claim of the hypothesis, that "learning" cannot turn into "acquisition," cannot be empirically tested. When a theory cannot be tested, it cannot be claimed sound (McLaughlin, 1987, p. 21, 56).

Input (Comprehension) Hypothesis

Closely in line with the acquisition-learning distinction hypothesis is Krashen's well-known input (or comprehension) hypothesis (1985, 2002, 2003). The input hypothesis is an effort to explain how language acquisition occurs. Krashen argued that learners acquire L2 when and only when they obtain "comprehensible input," i.e., when they understand messages (2003, p. 4). If their current L2 level is at stage *i*, they can

progress to the next stage, $i + 1$, only if they receive comprehensible input that contains the $i + 1$.

Criticisms. The input hypothesis has been criticized as being circular. Krashen claimed that an input is comprehensible because it promotes acquisition and that it promotes acquisition because it is comprehensible. McLaughlin (1987) pointed out that the idea is circular and does not lead anywhere. He argued that in order for a theory to be tested in a meaningful way, Krashen has to clearly define as well as demonstrate how to evaluate "comprehensible input" independently, which he has yet to do.

Probably due to lack of clear definition, Krashen's arguments sometimes contradict themselves. For example, in advocating the natural input approach, he claimed that teachers "need not know exactly where each student is in his or her developmental path; all we need to do is to provide a great deal of comprehensible input" (2003, p. 6). Nevertheless, it appears that if an input does not contain the student's $i + 1$ level, it cannot be claimed "comprehensible input."

In addition, Krashen apparently held double standards for form-focused instruction and natural input instruction. In the above argument favoring the natural input approach, he claimed that a teacher need not know where the student' current language level is. However, when speaking against form-focused instruction, his rationale is that teachers "usually guess wrong" at what the students' $i + 1$ is (2002, p. 220). An accurate evaluation of students' developmental level matters in one approach but does not matter in the other. The double standards diminish the credibility of his theory.

The input hypothesis has also been criticized by Schulz (1991), who disagreed with Krashen's (2003) view of comprehensible input as the "only" source for L2 acquisition (p.

4). Schulz (1991) reasoned that it is common personal experience that "skills which at one time were learned consciously through segmentation and analysis can eventually become automatic through practice and be available for spontaneous use" (p. 21).

In one of his recent publications, Krashen cited several empirical studies in attempt to claim the "astounding" "superiority" of natural input over skill-building (2003, p. 9). A careful examination of the data by the researcher of the present study, nonetheless, has revealed his "evidence" as problematic. By giving examples as to how Krashen has reported the results of the first three studies, the researcher of the present study is going to demonstrate his bias and some misleading interpretations.

In the first study (Asher, 1977), total physical response (TPR) approach was used as an example of a natural input approach in comparison to the traditional foreign language approach, namely students repeating after teachers, the grammar translation method, and reading and writing involving grammar instruction. The TPR approach involved students of Grades 5, 6, and 7-8; the traditional approach involved students of Grades 7-8 and 9. Krashen failed to make comprehensive and unbiased interpretations of the results. On several occasions he chose to report findings that have favored TPR but ignored those that have favored the traditional approach. For example, he pointed out that Grade 6 subjects receiving TPR outperformed Grade 9 subjects receiving the traditional approach. Yet he disregarded the fact that Grade 5 subjects receiving TPR did not do as well as Grade 9 subjects receiving the traditional approach.

Moreover, there are several reasons that comparisons across grade levels could be inappropriate. Prior instruction received by different grades could either positively or negatively impact the effects and thus contaminate the results. While one can argue that it

is easier for older students to digest the same materials than younger students, one can also reasonably argue that younger students could be more receptive to a foreign language. Either way it makes the comparisons across grade levels inconclusive and pointless. Now the only appropriate comparison left appears to be the one between Grade 7-8 TPR subjects and Grade 7-8 traditional-approach subjects. In both the listening and reading tests, the traditional-approach subjects outperformed the TPR subjects. Yet, Krashen refused to acknowledge the positive effect of the traditional approach, claiming the result was caused by longer hours of the language exposure that the traditional-approach subjects have received. Krashen's logic was simple: If TPR subjects outperformed the traditional subjects, it was due to the effectiveness of TPR; but if the traditional-approach subjects outperformed the TPR subjects, it was due to the traditional subjects' longer exposure to the language, not the merits of the approach. As a result, the traditional approach has been placed in a lose-lose position from the onset. The comparison, therefore, was virtually meaningless.

The second study Krashen cited in attempt to claim the "astounding" "superiority" of natural input over skill-building (2003, p. 9) is a comparison between the natural approach and the grammar translation method (Hammond, 1988). The results of the study do not support Krashen's claim in two ways: The probability level ($p < .07$) is over the limit of the conventional probability level of $p < .05$ and risks committing Type I error; the effect size is small ($d = .15$) and thus lacks practical implication.

The third is a comparative study between comprehensible input and traditional grammar audio-lingual methodology (Nicola, 1990). Although Krashen still argued for the "superiority" of comprehensible input, based on the standard of $p < .05$, the data have

indicated that in 9 pairs of comparisons, the majority (7 pairs) showed no significant difference between the two methods.

Studies that challenge the input hypothesis. Hammerley (1987) argued against Krashen's input hypothesis from yet another angle. Krashen indicated that if comprehensible input is "plentiful," the necessary grammar is automatically supplied (2003, p. 6), and grammatical accuracy will be obtained (2002, p. 220). To challenge Krashen's point of view, Hammerley (1987) reviewed six studies that investigated the effect of immersion programs that were based on the natural approach without explicit grammar instruction. It was found that students in such programs acquired very good listening and reading comprehension, but were "far from linguistically competent" (p. 395) in terms of speaking and writing production. The findings of the six research studies suggested that, in order to communicate freely beyond their limited linguistic capacity, students in immersion programs soon " 'fossilize' certain ungrammatical forms in their interlanguage, which mostly could be "terminal." Hammerley opposed the idea that the language classroom is a natural second language acquisition environment. "There is nothing natural about learning another language within four classroom walls," he argued (p. 398).

Similar conclusions were drawn from a separate study on immersion students in Canada. The subjects in the immersion program had ample opportunity to receive comprehensible input in the target language, yet they still committed some basic morpho-syntactic mistakes (Larsen-Freeman, 2003). Merely natural input without explicit instruction is apparently insufficient for L2 students to acquire basic production skills. In the learners' effort to communicate freely without proper grammar instruction, learners

"are forced to adopted or invent communication strategies that lead to fossilization" (Higgs & Clifford, 1982, p. 78).

According to White (1987), for L2 learners, it is in fact "incomprehensible" input, rather than comprehensible input as Krashen has proposed, that encourages the needed grammar development (p. 98). That is, when learners are unable to interpret the input in terms of their current grammar knowledge (or *i*), a restructuring of existing grammar takes place to make sense of the input. Restructuring was defined by McLaughlin (1990b) as "a new internal organization" as a learner moves from one level to the next in language development (p. 117).

Theories Emphasizing Consciousness-Raising

If second language acquisition theories are placed on a continuum, the theories that focus on natural input will be at one side of the continuum while those that emphasize consciousness-raising will be at the opposite side.

Attention and Second Language Acquisition

While Krashen (2002) argued that conscious learning does not turn into acquisition, numerous researchers have disagreed (e.g., Bialystok, 1978; Crookes, 1991; Schmidt, 1990, 1993, 1995). They contended that explicit instruction enhances language competence, and that practicing at a conscious level can lead to automaticity (i.e., acquisition).

Noticing hypothesis. In contrast to Krashen's input hypothesis, Schmidt (1990, 1993, 1995) claimed that learners must consciously notice forms and the meaning these forms realize in the input. "What learners notice in input is what becomes intake for learning" (1995, p. 20). He argued that attention to input is essential for input to become intake so that it will be available for further cognitive processing. This notion is shared by

Ellis (1993a), who suggested three possible outcomes when learners notice a certain grammar feature in input: (a) Learners notice, construe, and relatively instantaneously incorporate the new grammar feature to their interlanguage systems; (b) they notice, construe, and construct a conscious, explicit representation of the grammar feature; and (c) even if they forget the new grammar element after they notice and interpret, a "trace" will be stored and help them to process the grammar structure more thoroughly in the future (p. 75). Schmidt (1990) emphasized that it does not matter whether learners notice the grammar feature purposely or unintentionally. "If noticed, it becomes intake" (p. 139).

As Larsen-Freeman (2003) pointed out, the terminology regarding "noticing" requires clearer definition. Some scholars use "attention," "awareness," "consciousness," and "detection" interchangeably with "noticing" (see also McLaughlin, 1990a; Tomlin & Villa, 1994, p. 185). Nevertheless, although no agreement has been reached on how many types of attention there are, a review of literature conducted by Larsen-Freeman (2003) has shown that most SLA researchers agree on the importance of promoting noticing. It is also worth noting here that Schmidt (1990) explicitly ruled out the likelihood of "subconscious noticing" (p. 139).

Automaticity. In contrast to Krashen's dichotomy between learning and acquisition, many scholars view learning and acquisition as processes in one single continuum in which practice (i.e., learning) can lead to automaticity (i.e., acquisition). When there has been a habitual relationship between a particular type of input and some output pattern, one can claim that the process is "automatic" (Gass & Selinker,1994, p. 154). Bialystok (1978) indicated that information learned via formal instruction and stored in the explicit linguistic knowledge domain could be transformed into implicit linguistic knowledge

through practice. In a similar vein, Crookes (1991) contended that practice can lead to automaticity and suggested paying special attention to designing learning activities to determine what is to be extended to the domain of automaticity. As Sharwood Smith (1981) put it,

> Some aspects of second language performance can in principle be planned from the start entirely on the basis of explicit knowledge. . . . [I]t is surely reasonable to suppose that a certain number of structures planned and performed slowly and consciously can eventually develop into automatised behaviour. (p. 166)

These points of view seem to shed light on the value of explicit grammar instruction.

Research on Explicit Grammar Instruction for L2 Learners

Numerous researchers (e.g., De Graaff, 1997a, 1997b; Long, 1983; Norris & Ortega, 2000, 2001) have explored the effectiveness of explicit grammar instruction and concluded that explicit grammar teaching appears beneficial to L2 learners. For example, Long (1983) reviewed 11 studies that examined the effect of explicit "instruction" in comparison to natural "exposure." Among the 11 studies, six studies showed instruction as more productive, two were ambiguous, and three had null findings. Long concluded that the positive effect for explicit instruction holds for learners of different age (i.e., children and adults), for learners of different performance levels (i.e., beginning, intermediate, and advanced), on various types of tests (i.e., discrete-point and integrative tests), and in acquisition-rich as well as acquisition-poor environments.

More recently, Norris and Ortega (2001) scrutinized 77 experimental or quasi-experimental research studies published between 1980 and 1998. Among these studies, 49 that compared explicit L2 grammar instruction with least attention-focused or

pure exposure were included in a meta-analysis. Results of the meta-analysis found explicit types of instruction more effective than implicit types. A substantial average effect size was observed (d = .75). In addition, since the effect sizes for delayed posttests stayed relatively large, the effects of explicit grammar instruction appeared durable.

De Graaff (1997a, 1997b) examined the effect of explicit instruction on the acquisition of an artificial language and Spanish as foreign languages by adult native speakers of Dutch. Acquisition was tested in terms of accuracy and complexity of structures. The results of the study indicated that explicit knowledge of the target languages facilitates the acquisition of implicit knowledge of form and meaning of those languages. The findings of De Graaff's research, along with Long's (1983) and Norris and Ortega's (2001), have validated Schmidt's (1990, 1993, 1995) noticing hypothesis, which argues that noticing is necessary for input to become intake.

An Integrated Approach

As Dewey (1963) pointed out, human beings tend to think in terms of "either-ors" dichotomy and fail to realize the existence of intermediate possibilities. While the above review of literature seems to have provided reasonable evidence in support of the effectiveness of explicit instruction over teaching without explicit instruction, it by no means diminishes the value of natural input and communicative context in L2 learning. Whereas Krashen's (2003) view of comprehensible input as the "only" way to L2 acquisition (p. 4) does not hold, neither should the opposite assumption, that explicit instruction as the only way to acquisition, be held true. While some studies showed that context-based instruction without explicit grammar instruction has produced "linguistically faulty" output (e.g., Hammerley, 1987, p. 397), other studies showed that

explicit grammar instruction without context-based instruction has produced a "limited range" of communicative functions (e.g., Sorace, 1985, p.239). Canale and Swain (1980) recommended positioning grammatical competence within the scope of communicative competence. Savignon (1972, 1991), who coined the term "communicative competence" to illustrate the capability to communicate with other people in lieu of the ability to perform on discrete-point grammar tests (1972), also admitted that "involvement in communicative events . . . necessarily requires attention to form. Communication cannot take place in the absence of structure, or grammar" (1991, p. 268)

Isik (2000) compared the effects between a program of comprehension-based instruction supported by form-focused instruction and a basically form-focused program. He concluded that the comprehension-plus-form program was more effective than the form-focused-only instruction. Other researchers (Lightbown & Spada, 1990; Spada & Lightbown, 1993; Tomasello & Herron, 1989) also found that teachers who incorporated grammar instruction and corrective feedback into the context of communicative teaching were more effective than those who had never worked on grammar or who conducted only grammar instruction in isolation. Spada and Lightbown (1993) indicated that the effects were positive in both the short and long term.

Based on the above literature, an integrated approach seems warranted for L2 learners to maximize their learning. Many SLA researchers have advised the combinational use of form-focused and communicative approaches. Long (1991), Nassaji (2000), and Lightbown and Spada (1990) suggested teaching grammar within a comprehension-based or communicative approach to prevent learning in isolation as well as to advance learners' grammatical accuracy and overall communicative fluency. Ellis

(1993b) also urged structural grammar instruction "be used alongside some kind of meaning-based syllabus, which is designed to provide learners with opportunities for communicating in the second language" (p. 91). Since both the form-focused instruction and communicative approach have advantages, it appears reasonable to adopt a methodology encompassing both elements in an L2 classroom.

Pedagogical Practices for Grammar Teaching

If pedagogical practices for grammar teaching are placed on a continuum based on the level of consciousness-raising that is intended, they can range from the most explicit rule-articulation strategies to those with mere exposure to grammatical structure. This section presents various grammar pedagogical practices, some of which have been reviewed by Barnitz (1998) and Larsen-Freeman (2003). In general, the strategies can be divided into three major categories: input-oriented, output-oriented, and output-input combination.

Input-Oriented Strategies

There are two major types of input-oriented strategies: those that involve explicit grammar rule articulation, and those that promote noticing though exposure.

Explicit Grammar Rule Articulation

Explicit grammar rule articulation can be conducted in two ways. Teachers can present an explicit explanation of certain grammar points, or they can have the students figure out the targeted grammar features themselves (see, for example, Fotos and Ellis' task-based approach, 1991, p. 611).

Promoting Noticing through Exposure

To promote noticing through exposure, teachers can use reading materials as a vehicle. For example, they can try to attract students' attention to a certain grammatical structure by underlining, using boldface, or using different fonts in written texts. They can also use the "input flood" strategy by providing texts with high frequency of the target structure (Larsen-Freeman, 2003). In addition, Barnitz (1998) proposed reading and listening extensively to stories and musical lyrics, as well as participating in choral readings and Readers Theater. Another strategy that promotes noticing through exposure is Speaker and Speaker's (1991) "sentence collecting" strategy, which invites learners to collect sentences that are of interest to them in structure, function, or meaning. The collections take place in students' daily environments, either home or school, and the teacher exhibits them in the classroom to motivate reflections and conversations.

Output-Oriented Strategies

Swain (2005) argued that learners need to produce language output in order to discover the gap between what they are trying to say and their actual language competence. Recognizing the gap will in turn help learners become more receptive to future associated input. There are two major types of output-oriented strategies: those that involve sentence manipulation, and those that involve modeling language after authentic texts.

Sentence Manipulation

Learners can become aware of the array of sentence patterns and transformational processes of language by manipulating sentences (Barnitz, 1998). Examples of sentence manipulation for the purpose of grammar competence enhancement include sentence expansion, which allows students to reinforce complex sentence structures by adding

vocabulary, phrases, and clauses to a simple sentence (Barnitz, 1998; Cudd & Roberts, 1994), as well as sentence combining, which, as Barnitz (1998) pointed out, helps enhance learners' awareness of transformation-related processes, including embedding, permuting, coordinating, substituting, adding, and deleting. (For more discussion on sentence combining, see Enginarlar, 1994; Evans, Venetozzi, Bundrick, & McWilliams, 1988; Neville & Searls, 1991; O'Hare, 1973; Weaver, 1996).

Modeling Language after Authentic Literary Texts

Barnitz (1998) argued that syntactic competence can also be developed through interaction with authentic texts. He suggested using song lyrics or literature and encouraging learners to compose similar texts of their own. For example, he recommended using authentic texts such as *When I Was Young in the Mountains* (Rylant, 1982) and having students write their own memoirs using the same sentence structure. This strategy allows learners to develop structures associated with literature language. During the same time they are creating written products, they are also learning from the model of the authentic texts.

Output-Input Combination

The following pedagogical practices may facilitate grammar competence through a combination of output and input process.

Collaborative Dialogue

As its name suggests, collaborative dialogue requires students to work in pairs on a collaborative project. Pica's (1994) take on negotiation is in support of the use of collaborative dialogue: "As they negotiate, they work linguistically to achieve the needed comprehensibility, whether repeating a message verbatim, adjusting its syntax, changing

its words, or modifying its form and meaning in a host of other ways" (p. 494). According to Swain and Lapkin (1998), learners in collaborative dialogue, like all other learners, every now and then will stumble upon linguistic problems. Therefore, when they work in pairs they go through a cognitive process of language generation, peer assessment, and peer correction (p. 321, 333). It is very close to Slavin's (1995) model of cooperative learning in this regard.

Dialogue Journal

Using dialogue journals is a strategy advocated by Peyton and her colleagues (Peyton & Reed, 1990; Staton, Shuy, Peyton, & Reed, 1988). The goal of dialogue journals is to stimulate authentic exchange of language between learners and teachers so that the learners will develop functional and structural aspects of the target language in a meaning-making context (Barnitz, 1998; Orem, 2001). In the communication process, instead of pointing out the learners' grammatical errors, the teacher models the accurate forms so that the learners could "see them in the natural context of the dialogue" (Orem, 2001, p. 74). Modeling of appropriate language structure can be provided with questions or repetition of the students' ideas (Arey, 1993). A variation of the dialogue journal strategy is Ho's (2003) audiotaped dialogue journal, which extends the learner-teacher communication to the speaking dimension of the language arts. As Swain (2000) pointed out, collaborative dialogue, whether in writing or speaking, "is where language use and language learning can co-occur. It is language use mediating language learning. It is cognitive activity and it is social activity" (p. 97).

Writer Conference

Barnitz (1998) argued that syntactic skills are best obtained through authentic writing, reading, and editing processes. He recommended using writer conferences, in which the teacher, based on the students' writing pieces, raises questions about specific sentence meanings and structures in the context of natural communication (see also Au, 1993; Calkins, 1986; Weaver, 1990).

The Garden Path Technique

Another pedagogical choice that involves both input to and output from the students is the Garden Path technique (Tomasello & Herron, 1989). It is worth noticing that this grammar teaching pedagogy is not as open-ended as collaborative dialogue, dialogue journal, and writer conference. Instead, the teacher has an agenda from the onset and deliberately leads the students through a certain set of process. The Garden Path technique involves having students translate sentences that are easily misparsed or over-generalized, followed by the teacher's immediate feedback and correction. In other words, a learning situation is constructed in which students are expected to make errors in their language output so that the teacher can follow with input on grammar rule articulation in a systemic fashion. This technique allows the students to learn from their own mistakes and, according to Tomasello and Herron (1989), students who learn through the Garden Path output-input process tend to retain the grammar elements longer than those who merely have input instruction.

Summary on Pedagogical Practices for Grammar Teaching

This section has presented an array of grammar teaching pedagogies under three categories: input-oriented, output-oriented, and output-input combination. Table 11 provides a quick view of these pedagogies for the reader's reference.

Within each category some grammar teaching pedagogies provide more explicit rule articulation than others. The purpose here, however, is not to compare explicit with implicit instruction or to identify the single "best" pedagogical practice for grammar

Table 11
A Quick View of Pedagogical Practices for Grammar Teaching Reviewed

Categories	Pedagogical Practices for Grammar Teaching
Input-oriented	Explicit grammar articulation ■ Teacher providing explicit explanation of grammar features ■ Students figuring out grammar features in assigned texts themselves Promoting noticing through exposure ■ Using underlining, boldface, or different fonts in reading materials ■ Reading/listening extensively to stories ■ Input flood strategy ■ Sentence collecting strategy
Output-oriented	Sentence manipulation ■ Sentence expansion ■ Sentence combining Modeling language after authentic texts ■ Literature ■ Song lyrics
Output-input	Output-input combination ■ Collaborative dialogue ■ Dialogue journal ■ Audiotaped dialogue journal ■ Writer conference ■ Garden Path technique

teaching. As discussed earlier in the Second Language Acquisition and Second Language Learning section, it appears that an integration of explicit form instruction and communicative approach can best maximize L2 learning. The current section, therefore, provides a survey of grammar teaching strategies with various levels of explicit and implicit instruction that can be used in combination with one another based on the unique need in each individual classroom.

Chapter Conclusion

The review of literature on Taiwanese EFL instruction has indicated that there is a need to reform EFL education in Taiwan. Studies on the implementation of cooperative learning have presented the approach as a reasonable alternative to the traditional, whole-class approach in Taiwanese classrooms. Nevertheless, while several studies in Taiwanese settings have shed light on its potential benefits, the number of studies on the topic is still limited. More carefully structured studies need to be conducted to provide robust evidence on the motivational and cognitive effects of this approach. This study has been an effort to contribute to deeper understanding on effects of cooperative learning in EFL education in Taiwan.

The literature review conducted in this chapter has contributed to the research design of the present study. The review of literature on second language acquisition and learning has helped the researcher to realize the need to design teaching without having to fall into the "either-ors" dichotomy as Dewey (1963) has cautioned against. Various levels of explicit and implicit grammar learning activities, therefore, were included in the cooperative learning group as well as the whole-class instruction group to facilitate form-focused instruction in a contextualized and meaningful way. Furthermore, as a result

of reviewing a wide range of pedagogical practices for grammar teaching, the curricula for both the cooperative learning and whole-class groups were enriched with a variety of learning activities. They will be further discussed in the next chapter.

CHAPTER THREE

METHODOLOGY

The purpose of this study was to examine the effectiveness of Slavin's model of cooperative learning (1995, p. 45) on cognitive and motivational measures using a quasi-experimental research design in the context of an English as a Foreign Language (EFL) course being taken by college students in Taiwan. To accomplish this purpose, three major research questions and five exploratory questions were investigated. The main research questions are as follows:

1. How does motivation differ between the group receiving cooperative learning and the group receiving whole-class instruction?
2. How does utilization of learning strategies beyond class settings differ between the group receiving cooperative learning and the group receiving whole-class instruction?
3. How does grammar achievement differ between the group receiving cooperative learning and the group receiving whole-class instruction?

On account of the analysis on the above questions, the following five exploratory questions were investigated.

A. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on motivation? If so, what is the cause of the interaction?

B. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on out-of-class utilization of learning strategies? If so, what is the cause of the interaction?

C. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on grammar achievement? If so, what is the cause of the interaction?

D. How does student achievement differ between the cooperative learning group and the whole-class instruction group in terms of different cognitive levels?

E. What are the relationships among prior English level, gender, grammar achievement, task value, self-efficacy, use of elaboration strategies, and out-of-class peer collaboration behaviors?

To answer the three major research questions, it was hypothesized that (1) the group receiving cooperative learning in a college class in Taiwan would show higher motivation than the group receiving whole-class instruction; (2) the group receiving cooperative learning would show more frequent utilization of learning strategies beyond class settings than the group receiving whole-class instruction; and (3) the group receiving cooperative learning would have higher level of English grammar achievement than the group receiving whole-class instruction.

The null hypotheses for the three main research questions are as follows:

1. There is no statistically significant difference in the adjusted motivation means between the group receiving cooperative learning and the group receiving whole-class instruction.

2. There is no statistically significant difference in the adjusted means between the group receiving cooperative learning and the group receiving whole-class instruction in terms of use of learning strategies beyond class settings.
3. There is no statistically significant difference in the adjusted grammar achievement means between the group receiving cooperative learning and the group receiving whole-class instruction.

This chapter is organized around three sections. The first describes the subjects of the study in terms of the sampling procedure, sample size, sample characteristics, as well as external validity issues. The second discusses the three instruments being used in terms of their purposes, technical characteristics, scoring procedures, and score interpretation. The third describes procedures in terms of research design, data collection, and data analysis.

Subjects

The purpose of this section is to describe the subjects of the study in terms of the sampling procedure, sample size, personal and demographic characteristics, and external validity issues.

Sampling Method and Procedure

The researcher sought instructors in a specific university in Taiwan who volunteered to provide manipulated pedagogies based on the design of the study. One instructor was selected based on factors including education (master's degree or above), teaching experience (5 years minimum in EFL), teacher evaluations (3 points or above on a 5-point scale), professional training (regular participation), fluent command of English, as well as study and travel experience in an English-speaking country. Two freshman

classes of this instructor were then used for the study. One was randomly assigned as the control group and the other as the experimental group. The instructor had experience in implementing both whole-class instruction and cooperative learning but had not solely used cooperative learning throughout a semester prior to the present study. The instructor indicated that overall she had no preference between cooperative learning and whole-class instruction.

Sample Size

The general rule in quantitative research is to use the largest sample possible so that the subjects' scores on measured variables could be more representative of population scores. Nonetheless, the number of subjects that can participate is often limited due to feasibility and financial concerns. Researchers hence have developed rules of thumb for determining the minimum sample size needed for different research methods. According to Gall, Borg, and Gall (2003), the minimum sample size is 15 subjects in each compared group for experimental research. This study used a sample size of 84 with 42 in the experimental group and 42 in the control group. A typical enrollment in a Taiwanese EFL class was between 35 and 50.

Sample Characteristics

The subjects were students at a private university in central Taiwan. They ranked between 40% and 77% on the national joint technology college and university entrance examination. The demographic information of the sample will be described in more detail in Chapter Four. The island of Taiwan has a total area of 36,000 square kilometers, or 13,900 square miles, and a population of 22.72 million as of June 2005. The official language is Mandarin. Taiwanese and Hakka are the two major dialects in Taiwan

(Government Information Office [GIO], 2005). Taiwanese is spoken by about 70 percent of the people. Although English has been a required course in college for many years, it is still merely a foreign language in this island country. Students study the English language to pass the course so that they can obtain their academic degree, but English remains more or less out of people's daily lives. However, in order to help Taiwanese citizens meet the future challenges of globalization, the Taiwanese government has enacted the Challenge 2008 National Development Plan. One emphasis of this project is to enhance the ability to master foreign languages, especially English (GIO, 2004).

The university was located in the Taichung-Changhua Greater Metropolitan Area; this area had the fastest population increase in 2001 (GIO, 2003) and the second fastest population increase in 2002 among Taiwan's metropolitan areas. In December 2002, Taichung City was the third most populated area in Taiwan with 6,099 people per square kilometer (GIO, 2004). The university had approximately 400 full-time faculty members and 10,000 full-time students.

External Validity Issues

According to Creswell (2002), threats to external validity are threats that "reduce an experimental researcher's ability to generalize sample data to other persons, settings, and situations" (p. 324). Threats to external validity include (1) the lack of ability to generalize findings to different groups, such as other age, geographical, racial, social, or personality group, (2) the lack of ability to apply findings from one setting to another, and (3) the lack of ability to apply findings to past or future situations.

Nonrandom sampling was used for the study due to feasibility. Two existing college English grammar classes were selected using convenience sampling. There was

no concern for the first threat because subjects of these two classes were students of various social backgrounds from all over Taiwan. Both the control and experimental groups consisted of male and female students of typical college age. The researcher ultimately would generalize no further than that of the characteristics of the sample.

As to the second threat to external validity, the results are likely to be generalized to private colleges and universities in Taiwan that have similar entrance scores. Finally, to increase external validity pertaining to the third threat, the study took place during the regular semester for the duration of 12 weeks. The results of this study are likely to be generalized to regular semesters only, not during other situations such as summer mini-sessions.

Instrumentation

A questionnaire, a proficiency test, and an achievement test were used for the study. The Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) measured the subjects' motivation and use of learning strategies. An English grammar proficiency test was administered as pretest. A grammar achievement test was administered as posttest to measure the achievement on the content of 24 selected units covered in the duration of the study.

Motivated Strategies for Learning Questionnaire

The MSLQ (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) is an instrument designed to measure college students' learning motivation and their employment of learning strategies for a college course. The MSLQ is anchored in a general cognitive view of motivation and learning strategies, with the student characterized as an active processor of information whose beliefs and cognitions provide valuable insight to

instructional input (Pintrich, Smith, Garcia, & McKeachie, 1993). The theoretical framework that underlines the MSLQ is an adaptation of a general expectancy-value model of motivation (Pintrich & De Groot, 1990). (See Chapter Two for a review of expectancy-value theories.) The MSLQ contains two main sections: a motivation section and a learning strategies section. The structure of the MSLQ is presented in Table 12.

Table 12
Structure of the Motivated Strategies for Learning Questionnaire (MSLQ)

Section	Component	Scale
Motivation	Value	1. Intrinsic Goal Orientation
		2. Extrinsic Goal Orientation
		3. Task Value
	Expectancy	4. Control of Learning Beliefs
		5. Self-Efficacy for Learning & Performance
	Test Anxiety	6. Test Anxiety
Learning Strategies	Cognitive/Metacognitive Strategies	7. Rehearsal
		8. Elaboration
		9. Organization
		10. Critical Thinking
		11. Metacognitive Self-Regulation
	Resource Management Strategies	12. Time and Study Environment
		13. Effort Regulation
		14. Peer Learning
		15. Help Seeking

Motivation

The motivation section consists of 31 items. The section is divided into three components: (1) Value, (2) Expectancy, and (3) Test Anxiety. Each component is further divided into various subscales. The components and the scales are described in more detail as follows.

Value. This component measures students' goal orientations and value beliefs for a course. It contains three scales: intrinsic goal orientation, extrinsic goal orientation, and task value scales. Intrinsic goal orientation refers to the degree to which students perceive a learning task in the course as an end to itself rather than as a means to an end. In other words, the students participate in a learning task to challenge themselves, to satisfy their curiosity, or to master the task. On the other hand, extrinsic goal orientation refers to the degree to which the students perceive a learning task as a means to an end. The students' participation in the task is motivated by external factors such as competition, grades, rewards, or performance. Task value concerns the degree to which the students perceive the course material in terms of interest, significance, and usefulness. Higher task value leads to higher degree of participation in learning.

Expectancy. This component measures students' expectancy for success in a course and contains two scales: control of learning beliefs and self-efficacy for learning and performance. Control of learning beliefs concerns the degree to which the students believe that their efforts to study will bring about positive results. Self-efficacy for learning and performance concerns the students' performance expectations and their confidence level in mastering tasks.

Test anxiety. This component measures students' test anxiety in a course. Specifically, it measures the students' pessimistic thoughts and emotional aspects of anxiety that undermine performance on an exam.

Learning Strategies

The learning strategies section is divided into two components: (1) Cognitive and Metacognitive Strategies, and (2) Resource Management Strategies. Each component is

further divided into various scales. The components and the scales are described in more detail as follows.

Cognitive and metacognitive strategies. This component measures students' use of cognitive and metacognitive strategies with the following scales: rehearsal, elaboration, organization, critical thinking, and metacognitive self-regulation. Rehearsal refers to reciting items from the material and activating information in working memory. Elaboration refers to making associations between substances to be learned by paraphrasing, summarizing, and synthesizing. Organization refers to making connections between substances to be learned by selecting main ideas, outlining, making diagrams, or making tables. Critical thinking involves making evaluations and applying prior knowledge to new contexts for problem solving. Metacognitive self-regulation measures students' awareness and management of cognition by planning, monitoring, and regulating. These five scales consist of 31 items.

Resource management strategies. This component measures students' use of resource management strategies with the following scales: time and study environment, effort regulation, peer learning, and help seeking. Time and study environment refers to time management that not only requires planning and scheduling but also effective use of the time that was set aside for study. Environment management measures the degree to which students are able to manage and regulate their study environments. Effort regulation measures students' ability to commit to their learning goals by regulating their effort and attention when they face distractions, difficulties, or boring tasks. Peer learning measures students' use of peer collaboration strategy in learning (see also Pintrich, 2004). Help seeking measures students' ability to manage the supports of others by identifying

someone that is able to provide assistance and then actively seek for help. These four scales consist of 19 items.

Technical Characteristics

The development of the MSLQ began informally in 1982 and formally in 1986. Before the final version, the instrument went through several revisions to fit the conceptual model of the expectancy-value theory. These early versions were administered to college students in 1986, 1987, and 1988. After each data collection, the instrument was examined with statistical and psychometric analyses, including factor analyses, internal reliability coefficient computation, and correlations with academic performance and aptitude measures. After that, the items were rewritten and a new version constructed. This final version of the MSLQ was administered in 1990 to 356 Midwestern college students from thirty-seven classrooms, fourteen subject domains and five disciplines. The validity and reliability of the instrument are discussed in the following sections.

Validity. Given the nature of the underlying constructs being measured by the MSLQ, evidence of construct validity was needed. Confirmatory factor analyses with the lambda-ksi estimates were conducted by the authors of the instrument to test the construct validity of the MSLQ scales. According to Gall, Borg, and Gall (2003), factor analysis is a statistical technique that is used to reduce a large number of variables to a small number of factors by grouping moderately or highly correlated variables together. Values of .75 or higher for the lambda-ksi estimates show well-defined latent constructs.

Construct validity evidence of the 15 scales is presented in Appendixes A and B. Appendix A contains the items on Scales 1 to 6, which measure motivation orientation.

Appendix B contains the items on Scales 7 to 15, which measure the use of learning strategies.

Although several scales have estimates less than .75, given the fact that the sample (N = 356) spanned a broad range of courses from 14 subject domains and five disciplines (i.e., natural science, humanities, social science, computer science, and foreign language), the results from confirmatory factor analyses demonstrate reasonable construct validity for the MSLQ.

Reliability. Internal consistency was estimated with Cronbach alpha for each scale (Pintrich, Smith, Garcia, & McKeachie, 1993). Table 13 presents the reliability level for each MSLQ scale.

Table 13
Reliability Level for Each MSLQ Scale

	Scales	Cronbach Alpha
1	Intrinsic goal orientation	.74
2	Extrinsic goal orientation	.62
3	Task value	.90
4	Control of learning beliefs	.68
5	Self-efficacy for learning and performance	.93
6	Test anxiety	.80
7	Rehearsal	.69
8	Elaboration	.76
9	Organization	.64
10	Critical thinking	.80
11	Metacognitive self-regulation	.79
12	Time and study environment	.76
13	Effort regulation	.69
14	Peer learning	.76
15	Help seeking	.52

The Cronbach alphas of the motivation scales (Scales 1 through 6) ranged from .62 to .93; those for the learning strategies scales (Scales 7 through 15) ranged from .52 to .80. All scales are associated with adequate alpha reliability levels for the purpose of the study.

Questionnaire for the Present Study

Selection of scales. The self-report MSLQ was designed to be given in class. According to the constructors of the MSLQ, the 15 scales may be used collectively as an entire instrument or individually according to a researcher's or instructor's needs (Pintrich, Smith, Garcia, & McKeachie, 1991, p. 3; 1993, p. 804). Of the six scales in the motivation section of the MSLQ, the researcher selected the **self-efficacy for learning and performance** scale and the **task value** scale as measurement to narrow the focus of the study. Factors taken into consideration in the selection process include whether the scale has direct bearing to the research questions and whether it has reasonable reliability. The self-efficacy for learning and performance scale measures students' performance expectations and their confidence levels in undertaking tasks; the task value scale measures how learners perceive the course activities and materials in terms of their interest, importance, and relevance. Both are directed related to how well-motivated students could be in a course, and both have high reliabilities. The self-efficacy for learning and performance scale has a Cronbach alpha of .93, the higher between the two scales of the expectancy component; the task value scale has a Cronbach alpha of .91, the highest among the three scales of the value component.

Of the nine scales in the learning strategies section of the MSLQ, the researcher selected the **elaboration** scale and the **peer learning** scale as measurements for the study. Once again, factors taken into consideration in the selection include how directly the

scale is related to the research questions and how reasonable its reliability is. The elaboration scale measures learners' use of learning skills, such as paraphrasing, summarizing, and synthesizing; the peer learning scale measures their utilization of peer collaboration strategy in learning. Both have direct bearing to the research question on students' use of learning strategies; both are strongly related to Slavin's model of cooperative learning (1995, p. 45); and both have high reliability of .76. The four selected scales have a total of 23 items.

Modification and translation. For the present study, the four selected scales were given as pretest and posttest. For the motivation scales (i.e., the self-efficacy for learning and performance scale and the task value scale), all the items remained the same for the pretest and posttest. For the learning strategies scales (i.e., the elaboration scale and the peer learning scale), however, since the items could not evaluate a student's actual use of learning strategies in a course that had yet begun, the items were worded slightly differently for the pretest to assess a student's general use of learning strategies in college courses instead. See Table 14 for the 23-item questionnaire for the present study. When an item is worded slightly differently for the pretest, the revised part is bold-faced. The pretest and the posttest versions are divided by a slash (/) with the pretest version in the front. See Appendixes C and D for the complete forms of the pretest and posttest questionnaires with Likert-scaled options and instructions to the subjects. For the Chinese versions of the pretest and posttest questionnaires, see Appendixes E and F respectively.

Table 14
MSLQ Questionnaire Items Selected for the Study

Item
1. I think I will be able to use what I learn in this course in other courses.
2. It is important for me to learn the course material in this class.
3. I am very interested in the content area of this course.
4. I think the course material in this class is useful for me to learn.
5. I like the subject matter of this course.
6. Understanding the subject matter of this course is very important to me.
7. I believe I will receive an excellent grade in this class.
8. I am certain I can understand the most difficult material presented in the readings for this course.
9. I am confident I can understand the basic concepts taught in this course.
10. I am confident I can understand the most complex material presented by the instructor in this course.
11. I am confident I can do an excellent job on the assignments and tests in this course.
12. I expect to do well in this class.
13. I am certain I can master the skills being taught in this class.
14. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.
15. When I study for **a/this** class, I pull together information from different sources, such as lectures, readings, and discussions.
16. I try to relate ideas in **the subject of a course/this subject** to those in other courses whenever possible.
17. When reading for **a/this** class, I try to relate the material to what I already know.
18. When I study for **a/this** course, I write brief summaries of the main ideas from the readings and the concepts from the lectures.
19. I try to understand the material in **a/this** class by making connections between the readings and the concepts from the lectures.
20. I try to apply ideas from course readings in other class activities such as lecture and discussion.
21. When studying for **a/this** course, I often try to explain the material to a classmate or a friend.
22. I try to work with other students from **the same/this** class to complete the course assignment.
23. When studying for **a/this** course, I often set aside time to discuss the course material with a group of students from the class.

The Chinese version was translated by the researcher from its original English version. Care was taken to minimize the loss of original meanings through multiple examinations on the translations by two bilingual speakers of Mandarin Chinese and English, Drs. Yi-Guang Lin and Shengmei Chang. Dr. Yi-Guang Lin grew up in Taiwan, where he received his B.A. in Psychology from National Taiwan University in 1954. He received his Ph.D. in Psychology from the University of Michigan in 1962, and is currently an associate research scientist at the Center for Research on Learning and Teaching, University of Michigan. He has published 57 articles and book chapters in student learning strategies and motivation, self-concept, knowledge structure, test anxiety, student ratings of teaching, effectiveness of college teaching, and other related fields. Dr. Shengmei Chang was born in Taiwan, where she received her B.A. in Foreign Languages and Literature from National Chen Kong University in 1993. She received her Ph.D. in Curriculum and Instruction from the University of New Orleans in 2003 and is currently teaching Chinese at Defense Language Institute in Seaside, CA as an assistant professor.

In addition to the expert translation reviews, the Chinese questionnaires were given to three college students in Taiwan to (a) measure the administration time, (b) detect any ambiguity on the questionnaire, and (c) examine if the questionnaire was appropriate for Taiwanese educational contexts. Appendix G includes the permission letter that the researcher has obtained for use, modification, and translation of the MSLQ questionnaire.

Administration

The 23-item questionnaire was administered to both the control and experimental groups before and after the treatment. Each time the process took approximately 5

minutes. The students were ensured both orally and on the questionnaire that their answers would be confidential and had no bearing on the course grades.

Scoring Procedures and Score Interpretation

Items of the MSLQ were measured on a seven-point Likert scale with the first and the seventh points anchored as "not at all true of me" and "very true of me." To ensure that the subjects have a shared reference point when they responded to each item, the researcher had chosen to adapt slightly the instrument and had all response alternatives anchored. The adapted anchored response alternatives were "Never true of me," "Rarely true of me," "Occasionally true of me," "Sometimes true of me," "Often true of me," "Usually true of me," and "Always true of me." The Chinese translation of these terms were 完全不符合, 很不符合, 偶爾符合, 有時符合, 經常符合, 大致符合, and 完全符合. Because all of the items in the four scales selected for this study were positively worded (see Table 15) instead of negatively worded (see Table 16), no score needed to be reversed.

Table 15
Sample Item That Is Positively Worded

Getting a good grade in this class is the most satisfying thing for me right now.						
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Never true of me 完全不符合	Rarely true of me 很不符合	Occasionally true of me 偶爾符合	Sometimes true of me 有時符合	Often true of me 經常符合	Usually true of me 大致符合	Always true of me 完全符合

Table 16
Sample Item That Is Reverse Coded

I find it hard to stick to a study schedule. (REVERSED)						
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Never true of me 完全不符合	Rarely true of me 很少符合	Occasionally true of me 偶爾符合	Sometimes true of me 有時符合	Often true of me 經常符合	Usually true of me 大致符合	Always true of me 完全符合

After all the items were recorded, each scale was scored by averaging the item response scores. If there was any missing item for that scale, the average was calculated by dividing the sum of all the non-missing items by the total number of the non-missing items. The score of a scale was calculated only for subjects responding to at least 75% of the items on any given scale. The scoring resulted in each scale score ranging from 1.00 to 7.00. A score of 4.00 represented a neutral and moderate position.

For motivation scales, scores below 4.00 generally showed negative motivation in the course; scores above 4.00 showed positive motivation. For learning strategies scales, scores below 4.00 generally indicated less use of the learning strategy; scores above 4.00 generally showed more use of the learning strategy. Table 17 provides an interpretation of the various levels of score ranges.

Table 17
MSLQ Score Interpretation

Score	Response Alternative	Motivation	Use of Learning Strategies
1.00-1.50	Never true of me	None or exceptionally low	Never
1.51-2.50	Rarely true of me	Very low	Very infrequent
2.51-3.50	Occasionally true of me	Low	Infrequent
3.51-4.50	Sometimes true of me	Moderate	Moderate
4.51-5.50	Often true of me	High	Frequent
5.51-6.50	Usually true of me	Very high	Very frequent
6.51-7.00	Always true of me	Exceptionally high	Always

Reliability

After the 23-item questionnaire was administered to the students, results indicated that the questionnaire had a Cronbach alpha valued at .97, which gave the questionnaire a robust reliability. The individual Cronbach alphas for the four scales— self-efficacy for learning and performance, task value, elaboration, and peer learning—were also sound. They were found to be .97, .95, .93, and .70, respectively.

Grammar Pretest

Before the two groups of subjects received different teaching treatments, a 25-item pretest was administered to evaluate their grammar proficiency in a written test. These 25 items were selected from four forms of the General English Proficiency Test (GEPT), namely ER-0001P, ER-0002P, RTI-A, and RTI-B. The GEPT is a step test developed and administered by the Language Training and Testing Center (LTTC) in Taiwan. The test was initially commissioned in 1999 by Taiwan's Ministry of Education as part of its effort to promote foreign-language education and to offer a fair and reliable instrument to assess English proficiency.

The GEPT currently offers five levels of evaluation (basic, intermediate, high-intermediate, advanced, and superior) and includes four components (listening, reading, writing, and speaking) (Language Training and Testing Center [LTTC], n.d.). Forms ER-0001P and ER-0002P are basic-level reading tests; forms RTI-A and RTI-B are intermediate-level reading tests. Each of these forms has three sections: vocabulary and syntax, cloze-in-paragraph, and reading comprehension. The four tests contain 160 items totally.

Validity

The researcher scrutinized all 160 items of the four tests carefully and identified 25 items that measure grammar proficiency on the basic or intermediate level.

Reliability

The reliabilities of the four test banks and their individual sections were identified and presented in Table 18. All alphas showed reasonable reliabilities.

Table 18
Reliabilities of the Four Test Banks

	ER-0001P	ER-0002P	RTI-A	RTI-B
Section 1	.73	.77	.59	.47
Section 2	.69	.68	.49	.42
Section 3	.70	.78	.61	.48
Total	.87	.89	.79	.73

After the 25-item pretest was administered to the students, test results indicated that the test had a Cronbach alphas valued at .76, which gave the test a sound reliability.

Administration

The 25-item grammar pretest was given in class before the subjects received the treatment. It took approximately 30 minutes to administer.

Scoring Procedures and Score Interpretation

A subject got four points for each item answered correctly. Possible test scores ranged from 0 to 100. Since the test items were from the GEPT reading test banks, the GEPT standard was used as the guideline for interpretation of performance. In a GEPT reading test, a test taker needs to earn 67% of the total points to pass the section. For the

purpose of the study, therefore, a score of 67 (out of 100) was perceived the benchmark for satisfactory performance.

Grammar Posttest

A 162-item test bank was developed by the researcher for the purpose of measuring the English grammar achievement on the content of 24 selected units from the textbook *Grammar Express* (Fuchs & Bonner, 2001). The textbook was used by students of both the control and experimental groups. Features and activities of the book will be presented in the Research Design section under the subheading Teaching Material. The test items were developed around the content specified in the Table of Specifications (see Table 20) and addressed one of the six levels of the revision of Bloom's taxonomy of educational objectives (Anderson & Krathwohl, 2001). Each content area had at least one or more questions. The test items reflected what students had learned from teacher presentations and group or individual practice.

Validity

The achievement test covered six parts: present tense, past tense, present perfect tense and past perfect tense, future tense, wh- and tag questions, and modal verbs. Each part consisted of several units (e.g., Part I consisted of present progressive, simple present and non-action verb units). The units selected from each of these carefully defined domains reflect the researcher's judgment of the importance of any given unit. That is, the criteria of selection included the level of difficulty and significance. For example, for Part I (see Table 19), Units 4 and 5 were left out. Unit 4 was left out because it was basically a review unit for Units 1 and 2 and provided no new material to be learned. It made no sense to have test items from units that covered the same material. Unit 5 was

not covered because the content was too easy and would not be able to differentiate the learners' achievement levels. Thus, for Part I of the textbook, three units (Units 1, 2, and 3) were selected to be taught and to be covered in the achievement test.

Table 19
Sample Table of Contents from the Textbook

PART I	Present Tense	
	Unit 1	Present Progressive
	Unit 2	Simple Present Tense
	Unit 3	Non-Action Verbs
	Unit 4	Present Progressive and Simple Present Tense
	Unit 5	Imperative

Using this approach, the researcher chose to cover three units for Parts I and V, four units for Parts II, III, and IV, and five units for Part VI. Twenty-six test items were written for Part I, 25 items for Part II, 24 items for Part III, 29 for Part IV, 20 for Part V, and 38 for Part VI. It was obvious that Parts I, II, III and IV had approximately the same number of test items, each accounting for about 16% of the total items. Part V had the fewest items, accounting for only 12% of the total number. And Part VI contained the most items, accounting for 23% of the total items.

The test items were written for one of the six levels of cognitive processes: "remember," "understand," "apply," "analyze," "evaluate," and "create." These six levels of cognitive processes were based on Anderson and Krathwohl's (2001) revision of Bloom's taxonomy of educational objectives. It would be easier to understand the meanings of the cognitive processes if one reads the phrase "The student is able to . . ." or "The student learns to . . ." before the verbs mentioned above (p. 66). More specifically, a test item on the level of "remember" required a student to be able to retrieve pertinent

information. A test item on the level of "understand" required a student to be able to "construct meaning from instructional messages" (p. 70). A test item on the "apply" level entailed a student to be able to use the acquired knowledge to a new situation. A question on the "analyze" level entailed a student to be able to break information into its components and decide how the components associate with one another and with the whole structure. A question on the "evaluate" level required a student to be able to check and critique information based on imposed criteria. Finally, a test item on the level of "create" entailed a student to be able to "put elements together to form a coherent or functional whole" (p. 68).

In the process of developing the achievement test bank, the researcher realized that most of the content called for cognitive processes beyond the level of simply retrieving information. There were, therefore, fewer test items on the level of "remember." The test items on the level of "analyze" were divided into two types of questions. The first type required a student to break the text into parts and examine their associations so that the student could discern the correct answer for the missing part. The second also required a student to break the text into parts and analyze the relationships, but, in this case, it was necessary for them to find the incorrect part out of the whole structure. The first type of questions accounted for 29% and the second type 26% of the total test items. The percentage of test items for the six levels of cognitive processes were 2%, 7%, 10%, 55%, 14%, and 12%, respectively. The number of test items for each unit, part, and level of cognitive process is presented in Table 20.

Table 20
Table of Specifications for the Grammar Achievement Test Bank

Part	Unit	Content	Week	Level of Cognitive Processes							Chapter Total	Part Total
				Remember	Understand	Apply	Analyze (correct)	Analyze (incorrect)	Evaluate	Create		
I	1	Present Progressive	1	0	1	2	2	2	1	1	9	
	2	Simple Present Tense	1, 2	0	1	2	2	2	2	1	10	
	3	Non-Action Verbs	2	1	0	0	2	2	1	1	7	26
II	6	Simple Past Tense: Affirmative	3	1	1	1	2	2	0	1	8	
	7	Simple Past: Negative/Q	3	0	0	1	1	1	1	0	4	
	9	Past Progressive	4	0	1	2	1	1	1	1	7	
	10	Past Progressive/Simple Past	4	0	0	0	2	2	2	0	6	25
III	11	Present Perfect: Since/For	5	0	0	1	2	1	1	1	6	
	12	Present Perfect: Already/Yet	5	0	1	0	2	1	1	1	6	
	14	Present Perfect/Simple Past	6	0	0	0	2	1	1	1	5	
	17	Past Perfect	6	0	1	1	2	1	1	1	7	24
IV	19	Future: Be Going to and Will	7	0	1	1	2	1	0	1	6	
	20	Future: Contrast	7	0	0	0	2	2	1	0	5	
	21	Future Time Clauses	8	0	0	0	4	4	1	1	10	
	22	Future Progressive	8	0	1	1	2	2	1	1	8	29
V	24	Wh- Questions	9	0	0	0	2	2	1	1	6	
	25	Tag Questions	9	0	0	0	2	3	2	1	8	
	26	Additions	10	0	0	0	2	2	1	1	6	20
VI	27	Ability	10	1	1	1	2	2	1	1	9	
	28	Permission	11	0	1	1	3	3	1	1	10	
	29	Request	11	0	1	1	3	2	1	1	9	
	30	Advice	12	0	0	1	2	1	0	0	4	
	31	Suggestion	12	0	1	1	1	2	0	1	6	38
Total		24 units	12 weeks	3	12	17	47	42	22	19	162	162

Item Selection

The test bank appeared to contain too many items for the pilot study. The test's reliability could be distorted by the fatigue effect and violation of independence of items. The fatigue effect is declined performance on the dependent measure attributable to being exhausted when a test is too long (Mitchell & Jolley, 2004). Violation of independence of items could occur when test takers are able to answer an item correctly, not because they possess the knowledge, but because the answer to the item is given away by something contained in another item (Adkins, 1974; Mislevy, Steinberg, & Almond, 2002). Due to the above considerations, it was decided that approximately 50 items would be selected from the 162-item test bank. Because there were six carefully defined domains for the achievement test, the general guide was eight items per part, with at least two items per unit.

The researcher examined the six levels of cognitive processes. As discussed in the validity section, the percentages of test items for the six levels of cognitive processes (i.e., remember, understand, apply, analyze, evaluate, and create) are 2%, 7%, 10%, 55%, 14%, and 12%, respectively. On the "analyze" level, there were two types of items. Their percentages were 29% and 26%, respectively. The number of test items for each level was then determined by multiplying the above-mentioned percentages by 50 (the desired total number), which resulted in 1, 4, 5, 27, 7, and 6 items for each level, with 14 items for the type-one "analyze" and 13 items for the type-two "analyze." Because there was only one item on the "remember" level, it was then decided that all three items from the test bank would be kept for the pilot study, which resulted 52 total test items. Table 21

shows the number of test items in the test bank and the number that was selected for each level of cognitive process.

Table 21
Number of Test Items on Each Level of Cognitive Processes: A Comparison Chart

Level of Cognitive Processes	Remember	Understand	Apply	Analyze (Choose correct answer)	Analyze (Choose incorrect answer)	Evaluate	Create	Total
Number of Items in Test Bank	3	12	17	47	42	22	19	162
Percentage of Items in Test Bank	2%	7%	10%	29%	26%	14%	12%	100%
Number of Items to Be Selected for the Pilot Study	3	4	5	14	13	7	6	52
Percentage of Items to Be Selected for the Pilot Study	6%	8%	10%	27%	25%	13%	11%	100%

With the total number of items for each cognitive level determined, the researcher established the number of items needed for each unit and each part. Fifty-two (i.e., the total number needed for the pilot study) was divided by 24 (i.e., the number of units), and the result indicated that each unit needed at least two items. Subsequently, 52 was divided by 6 (i.e., the number of parts), and the result indicated that each part needed eight to nine items. Because Part VI contained five units, according to the criterion of the two-item minimum, this part needed at least 10 test items. Because Part I and Part V had fewer items, it was decided that these two parts would have eight and seven items each, while Part II, III, IV, and VI would have nine items each. The total number of test items for the pilot test would be 52.

The Table of Specifications for the Grammar Achievement Test Bank (i.e., Table 20) was then used to aid the process of item selection. As previously discussed, all three items for the "remember" level in the test bank were kept, so there was one item each on the "remember" level for Part I (Unit 3), Part II (Unit 6), and Part VI (Unit 27).

Next, it was decided that the items on the "create" level should be determined before those on the other levels to avoid violation of independence of items (Adkins, 1974; Mislevy, Steinberg, & Almond, 2002). Since there had been items selected from Part I, II, and VI, items on the "create" level of these parts were not considered. Because six items were needed on this level, it appeared that the most appropriate cluster of items would be the three items from Part IV (Units 19, 21, and 22) and Part V (Units 24, 25, and 26). Since the items on the "understand," "apply," and "evaluate" levels might become an aid for the test takers to complete the task on the "create" level, items on these three levels in the parts covered by the selected "create" level items were excluded from consideration.

The items on the "understand" level were considered next. Four items were needed based on Table 21. Because the items in Parts IV and V had been excluded, there were Parts I, II, III and VI that the four items could be selected from. It appeared reasonable to select one item out of each part. The first unit in Part I was Unit 1, so the item was selected from this unit. The first unit in Part II was Unit 6, but one item on the "remember" level had been selected, so the item from the next available unit, Unit 9, was selected. The same rationale applied to the selection of the rest of the items.

On the "apply" level, five items were needed out of four parts (not all six parts so that violation of independence of items could be avoided). After one item was selected

from each of the four parts, there was one more item to be selected. Because Part VI contained the most items on this level in comparison to the other three parts, that extra one item was selected out of Part VI. In addition, when there were more than one item in the same cell (i.e., same unit and same cognitive level), only one was selected. Other than these two situations, the same rationale above applied here for the item selection.

The next level of cognitive processes, the "analyze" level, contained the most number of items to be selected. Therefore, the selection process skipped this level of the selection until the items on the "evaluate" level were selected. The items on the "evaluate" level, like those on the "create" level, were constructed in clusters. The selection process began from the top of the table and the cluster of the four items in Part I was selected. Because seven items were required for this level, only a cluster with three items would meet the selection criterion. The cluster of three items in Part VI was selected.

Finally, there was only one cognitive level left—the "analyze" level. It was mentioned earlier that this level of items contained two item types: one required the student to choose the correct answer and the other required the student to choose the incorrect answer. In addition to the criteria that had been applied earlier in the item selection process, the selection on this level paid special attention to ensure that there were at least two items from each unit and that the total number of items in each unit meet the criterion. Furthermore, care was taken so that the two types of the "analyze" items were distributed as evenly as possible within and among units. Table 22 presents the results of the item selection.

Table 22
Table of Specifications for the Grammar Achievement Pilot Study

Part	Unit	Content	Week	Level of Cognitive Processes							Chapter Total	Part Total
				Remember	Understand	Apply	Analyze (correct)	Analyze (incorrect)	Evaluate	Create		
I	1	Present Progressive	1	0	1	0	1	0	1	0	3	
	2	Simple Present Tense	1, 2	0	0	1	0	0	2	0	3	
	3	Non-Action Verbs	2	1	0	0	0	0	1	0	2	8
II	6	Simple Past Tense: Affirmative	3	1	0	0	1	1	0	0	3	
	7	Simple Past: Negative/Q	3	0	0	1	1	0	0	0	2	
	9	Past Progressive	4	0	1	0	0	1	0	0	2	
	10	Past Progressive/Simple Past	4	0	0	0	1	1	0	0	2	9
III	11	Present Perfect: Since/For	5	0	0	0	1	1	0	0	2	
	12	Present Perfect: Already/Yet	5	0	1	0	1	0	0	0	2	
	14	Present Perfect/Simple Past	6	0	0	0	1	1	0	0	2	
	17	Past Perfect	6	0	0	1	1	1	0	0	3	9
IV	19	Future: Be Going to and Will	7	0	0	0	0	1	0	1	2	
	20	Future: Contrast	7	0	0	0	1	1	0	0	2	
	21	Future Time Clauses	8	0	0	0	1	1	0	1	3	
	22	Future Progressive	8	0	0	0	1	0	0	1	2	9
V	24	Wh- Questions	9	0	0	0	1	0	0	1	2	
	25	Tag Questions	9	0	0	0	0	1	0	1	2	
	26	Additions	10	0	0	0	1	1	0	1	3	7
VI	27	Ability	10	1	0	0	0	0	1	0	2	
	28	Permission	11	0	1	0	0	0	1	0	2	
	29	Request	11	0	0	1	0	0	1	0	2	
	30	Advice	12	0	0	1	0	1	0	0	2	
	31	Suggestion	12	0	0	0	1	1	0	0	2	10
Total		24 units	12 weeks	3	4	5	14	13	7	6	52	52

Pilot Test and Item Analysis

The researcher piloted the test on 34 Taiwanese students, scored the tests, and conducted a classical item analysis on the students' responses. The purpose of the analysis was multi-dimensional. First it was to determine whether each of these items functioned properly, second whether the level of difficulty was appropriate, third whether an item was able to distinguish high-scoring students from low-scoring students, fourth whether the keyed answers were accurate, and fifth, in the selective-response items (i.e., Items 1 to 39), whether each distracter was functioning as designed. For the selective response items, the analysis looked into four types of statistics: proportions of students choosing each response, the item facility index (also called the item difficulty index), response-total correlation, and the discrimination index (Ebel & Frisbie, 1991). For supply items (i.e., Items 40 to 52), the analysis examined the means and the discrimination index.

By calculating the proportion of the students choosing each response, two statistics were obtained: (a) the proportion of students choosing the correct response, and (b) the proportion of the students choosing each incorrect response. The proportion of students choosing the correct response was the item facility index, or the difficulty index. In this study the researcher has chosen to refer to the index as the item facility index because this index actually refers to the degree of *easiness* instead of the degree of *difficulty*. The maximum value of the index is +1.00, and the minimum value is 0.00. The closer the value is to +1, the easier that item is. For example, when the facility index was .75, 75% of the students answered an item correctly and the item was relatively easy.

When the facility index was close to 0.00, it indicated that almost no student answered the question right and that the item was not easy at all.

The proportion of the students choosing each incorrect response has yielded information as valuable as the facility index. By looking into these statistics, the researcher was able to tell how the distracters functioned. When a proportion was extreme (either too high or too low), the researcher would take a look at the item, including the stem and the responses to determine whether a distracter was attractive enough or too attractive and whether the stem was ambiguous in any way. In some cases a revision of an item was necessary to enhance the quality of the item.

The point-biserial correlation coefficient (r_{pbi}) was used to estimate item discrimination. The r_{pbi} is the correlation coefficient between score on a test item and score on the total test for each student (Ebel & Frisbie, 1991). When the researcher saw a high positive r_{pbi}, she knew that high-scoring students on the total test tended to get right on that specific item and low-scoring students tended to get wrong. High negative r_{pbi} showed the opposite relationship. That is, high-scoring students on the total test tended to get the item wrong and low-scoring students on the test tended to get it right. A low value of r_{pbi} showed little or no relationship.

An r_{pbi} was also used to estimate response-total correlations, which pointed out how the students choosing a specific response performed on the total test. For example, on Item 1, the correlation between response A and the total score was -.30. The negative value indicated that the students choosing this response tended to perform poorly on the total test. This was appropriate given that A was an incorrect response. The magnitude of

this index indicated the distracter functioned quite well. It attracted students who performed poorly and not attracting those who did well.

Appendix H shows the results of the item analysis. For the selective-response items (i.e., Items 1 to 39), statistics have included the proportions of students choosing each response, the facility index, the response-total correlations, and the r_{pbi} discrimination index. For the supply items (i.e., Items 40 to 52), means and the r_{pbi} discrimination index are presented. The researcher used these statistics to improve the quality of the items.

Item Evaluation through Cross-Examining the Results of the Item Analysis

After the above-described procedure of item analysis, the researcher obtained five statistics: the proportions of students choosing each response, the facility index, the response-total correlations, the r_{pbi} discrimination index, and mean scores for the supply items. The interplay of these statistics yielded important information on whether an item needed revision, and if so, how to improve the quality of that item. The following section discusses a sampling of this process. Table 23 is an excerpt of Appendix H. It shows the results of the item analysis for Items 1 to 3.

Table 23
Excerpt of the Item Analysis on the Pilot Test for the Grammar Posttest

Item	Proportion of Students Choosing Each Response				Item Facility Index	Response-Total Correlations				r_{pbi} Discrimination Index
	A	B	C	D		A	B	C	D	
1	.24	.41*	.27	.09	.41	-.30	.17*	.12	-.03	.17
2	.12	.29	.44*	.15	.44	-.21	-.26	.38*	-.01	.38
3	.18	.09*	.21	.53	.09	-.45	.48*	.08	-.01	.48

*Note. * denotes correct answer.*

In Item 1, the r_{pbi} discrimination index fell below the criteria of .19 (r_{pbi} = .17). According to Ebel and Frisbie (1991), a discrimination index below .19 indicates a poor item that might need to be rejected or revised. The researcher took a careful look into the other statistics. It turned out that the response-total correlations indicated a weak but positive relationship between the correct response, B, and the total test (r_{pbi} = .17), a negative relationship between both distracters A and D and the total test (r_{pbi} = -.30 and -.03, respectively), and little relationship between distracter C and the total test. Moreover, the item facility (*IF*) index shows a reasonable facility level (*IF* =.41). Generally this information indicates responses are attracting satisfactory proportion of students. The interplay of the statistics indicated that, although the r_{pbi} discrimination index was low, the overall evaluation indicated that it was a reasonable item.

In Item 2, the item facility index showed a reasonable facility level (*IF* = .44), each response functioned by attracting a satisfactory proportion of students, and the r_{pbi} discrimination index showed good discriminating ability of r_{pbi} = .38. In addition, while the correct response was positively associated with the total score, all the distracters were associated with the total score negatively (r_{pbi} = -.21, -.26, and -.01, respectively). Therefore, the examinations of these four statistics indicated a well-constructed item that required no revision.

On the other hand, if one only looked into the response-total correlations and the r_{pbi} discrimination index in Item 3, it seemed the item was well constructed. It had a strong discrimination index of r_{pbi} = .48; the correct response was positively and strongly associated with the total score; the distracters were either negatively associated or had almost no relationship with the total score. But if one also examined the item facility

index and the proportion of students choosing each response, one would realize that the facility level of this item was relatively low (*IF* = .09). While distracter D attracted more than half of the students, the correct response, B, only attracted approximately one tenth of the students to choose. Although the response-total correlation of r_{pbi} = .48 between response B and the total score indicated that the one-tenth of students who chose B tended to be high-performing students, this proportion was still low and called for attention to the construction of the item. Thus the researcher went back and examined the way this item was constructed. The examination showed that the item was somewhat ambiguous, and the researcher edited the stem to enhance the clarity of the item.

By now it should be clear to the reader that this was a process that called for interaction of the four statistics for each item. Not a single set of data could alone determine the effectiveness of an item. As illustrated in the analysis of Item 1, an insufficient discrimination index caught the researcher's attention, but a further examination of the other three sets of data indicated no need for revision. In contrast, Item 3 had good discriminating level, yet both the *IF* index and the proportion of each response chosen indicated need for revision. The interplay of these four statistics was very important. The rest of the selective-response items were scrutinized through the same careful procedure. Out of the 39 selective-response items (i.e., Items 1 to 39), nine items were revised.

The researcher then examined the 13 supply items (i.e., Items 40 to 52) with the same level of scrutiny by looking into the mean and the r_{pbi} discrimination index of each individual item. The mean indicated the performance of an average test taker on a specific item. It also showed the percentage of students who answered the item correctly.

The r_{pbi} discrimination index, once again, showed how well an item was able to distinguish high-scoring students from low-scoring students. The researcher paid special attention to items with a low mean score and/or a low r_{pbi} discrimination index. As a result, two supply items and the instructions for Items 40 to 46 were revised to adjust the levels of difficulty and to enhance clarity.

Administration

The 52-item posttest was given in class after the subjects had received the treatment. It took 50 minutes to administer.

Reliability

After the subjects took the 52-item grammar achievement posttest, results indicated that the test had a Cronbach alpha valued at .91, which gave the test a robust reliability.

Scoring Procedures and Score Interpretation

A student first received one point for each item answered correctly. To enhance objectivity and thus reliability of the scoring on the supply items, the following procedures (Ebel & Frisbie, 1991) were followed: (a) Answers were scored question by question instead of student by student, and (b) the identity of the student whose answer the scorer was scoring was concealed. The raw scores were then transferred into percentages answered correctly, i.e., Posttest Score = Raw Score * 100 / 52. Possible posttest scores ranged from 0 to 100. In order to have a shared reference point with the pretest, the posttest also adopted the GEPT standard for score interpretation, i.e., a score of 67 (out of 100) was perceived the benchmark for satisfactory performance.

Procedures

The section describes procedures in terms of research design, data collection, and data analysis. The procedures were reviewed and approved by the University Committee for the Protection of Human Subjects in Research, University of New Orleans.

Research Design

A quasi-experimental pretest-posttest comparison group research design has been chosen for the study to compare the cooperative learning group with the whole-class instruction group in terms of motivational and cognitive outcomes. The quasi-experimental design was selected due to the availability of the subjects. When developing artificial groups is unfeasible and intact classes accessible, the quasi-experimental design appears to be a reasonable choice (Creswell, 2002). The design is represented below:

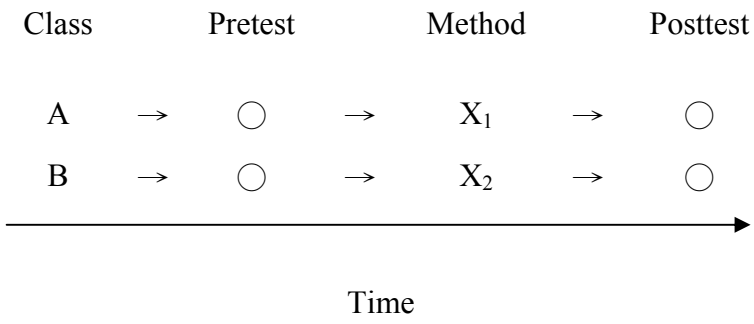

Figure 4. Quasi-experimental pretest-posttest comparison group research design.

The subjects in the study were students in two pre-existing university freshman English grammar classes. The researcher manipulated the types of instruction: One class was the control group receiving whole-class instruction; the other was the experimental group

receiving cooperative learning pedagogy. The treatment lasted for the duration of 12 weeks. Each class had 42 students. Both groups were administered a selective version of the Motivated Strategies for Learning Questionnaire (MSLQ) as a pretest and posttest, a pretest on English grammar proficiency, and a posttest on English grammar achievement. The two groups were compared in terms of their performances on the grammar tests, motivation, and out-of-class use of learning strategies.

Teaching Material

The main teaching material for both the control and experimental groups was a grammar book called *Grammar Express* (Fuchs & Bonner, 2001). Twenty-four units were covered during the 12-week period. The book featured illustrations, charts, notes, and exercises. Each unit began with an illustration that presented and contextualized the grammar point through a visual aid, which could be a cartoon, a photo with speech bubbles, an advertisement, or comic strips. The grammar point was also presented in charts that showed its various forms. Next, the grammar notes provided ample grammar explanations and example sentences. Graphics, such as timeline for verb tenses, often accompanied narrative explanations for easier understanding and better retention.

There were various exercise types that allowed learners to practice the grammar point in context. The exercises were mostly embedded in high-interest authentic texts, such as diaries, humorous conversation, and appealing news reports. Interesting photos or illustrations that aided the understanding of the texts were presented throughout the exercises. Finally, the review section allowed students to demonstrate their mastery of the structure through review and editing exercises.

As a result of reviewing a wide scope of grammar teaching pedagogies, the researcher designed curricula that allowed students in both the control and experimental groups to be exposed to context-rich activities, including communication activities and writing activities as recommended in the teacher's guide (Fuchs, 2002) to the textbook. For example, when working on the present progressive tense in Unit 1, the instructor brought in several magazine and newspaper photos and wrote some time expressions on the board, such as "right now," "at the moment," and "these days." Next, the instructor encouraged students to "report" what was happening on the photos using the time expressions on the board (p. 1). When working on non-action verbs in Unit 3, students wrote an end-of-the-day journal entry about how they felt. They were prompted to use non-action verbs that they had learned in that given unit, such as "feel," "want," "prefer," "know," "need," and "think" (p. 6).

Although the authors of the textbook (Fuchs & Bonner, 2001) and the teacher's manual (Fuchs, 2002) did not specify, the researcher of the current study scrutinized the content of each unit during the textbook selection phase and was able to identify a theme for most of the units. The origins of thematic units can be traced back to Dewey's (1963) principles on curriculum continuity and integration, and to Bruner's (1960) ideas on a spiral curriculum where learners can visit and revisit key concepts. In the context of the current study, the thematic units could help students learn the content in a context-rich, meaningful, and coherent manner. Table 24 presents the themes that were covered during the 12-weeks' study.

Table 24
Weekly Themes of the Grammar Lessons for Both the Experimental and Control Groups

Week	Theme(s)	Week	Theme(s)
1	Personalities	7	Classroom of the Future/Travel Plan
2	Feelings	8	Goal Planning/Robots
3	Poets/Interviews	9	Cross-Examination/Our City
4	Eye Witnesses	10	My Partner and I
5	Celebrities	11	Making Requests
6	Commuter Marriage/My Life	12	Jobs/Travel

In a nutshell, although the textbook followed a structural syllabus, in the meanwhile it provided ample opportunity for learners to have authentic use of English grammar in discourse. It was the intention of the researcher to use the textbook as a vehicle to provide learners in both the experimental and control groups form-focused instruction in a contextualized and meaningful way.

Independent Variable

There were two levels of independent variable for the study. They are the two types of instruction. The treatment variable for the experimental group was Student Teams Achievement Divisions (STAD); the treatment variable for the control group was whole-class instruction. Both types of instruction used the same textbook and covered the same material.

Procedure for the control group. As discussed in Chapter Two, the most usual way of teaching English in Taiwan has been the whole-class, teacher-centered grammar-translation method (Su, 2003) which consists of teaching of grammatical rules and sentence structures of English using Chinese translations. The instructional design for the control group in this study included the traditional whole-class grammar translation method, but in order to enhance the quality of teaching in the control group,

communication-based class activities were added to the curriculum. For example, before getting into the grammar point in each unit, the teacher used warm-up questions and visual aids to focus the students' attention and to set the context for the grammar point; the teacher invited the students to participate in answering those questions in the whole-class setting to activate their schemata; context-rich activities such as journal writing and editing were also included. But in order to differentiate the control group from the experimental group, which utilized peer learning, these activities were carried out either whole-class (e.g., whole-class discussion) or individually (e.g., journal writing). The control group used the same teaching material as the experimental group, including the textbook that has been introduced above.

Procedure for the experimental group. There are many ways cooperative learning can be implemented. The specific cooperative learning method used in this study was Student Teams Achievement Divisions (STAD). The STAD procedure for the experimental group was composed of five main steps: (1) instruction, (2) teamwork, (3) individual tests, (4) individual improvement scores, and (5) group average improvement points and team recognition. After the instructor presented her teaching, the groups were set to work. (The grouping procedure will be discussed at the end of the section before the discussion of treatment fidelity.) The students went over the same exercise materials. But instead of working individually, they worked together with their teammates. They helped each other answer and understand the materials through elaborated explanations, peer modeling, peer practice, and peer assessment and correction. It should be noted that based on the resource interdependence theory (Johnson, Johnson, & Smith, 1991), teammates often shared the materials instead of having their own copies (see Chapter

Two for details). When there were communicative activities, such as journal writing or topic discussion, the activities were group-based, in contrast to the control group's whole-class or individual approach.

Next, the students took tests individually. One's individual test score (ITS) was compared to one's pretest score, and the difference between ITS and the pretest score was one's "individual improvement score" (IIS). The IIS transferred to "individual improvement point" (IIP) (see Table 25). After the "group average improvement point" (GAIP) was calculated (see Figure 5 for the worksheet), the team accomplishments were recognized via the GAIP (see Table 26).

The instructor made it explicit from the onset that she welcomed every group to earn an award and that there would be no limit to the number of groups receiving each award. So students understood that they were not competing with other groups. They would just be working toward group achievement by challenging their personal past performance. Ten percent of a student's semester grade was based on the GAIP. Any points exceeding 10 were considered as bonus points toward a students' semester grade. In addition to IIP and GAIP scores, a student's semester grade was also determined by peer evaluation (by teammates) and cooperative process (as evaluated by the instructor). The comparison chart for the semester grade rubrics of the experimental and the control groups is presented in Appendix I.

Before cooperative work began, the students in the experimental group were sorted into 10 heterogeneous groups of four to five group members based on the pretest scores. Care was taken to ensure that each group consisted of learners whose achievement

Table 25
Conversion Table for IIS and IIP

IIS					IIP
		IIS	≤	-10	0
-10	<	IIS	≤	0	10
0	<	IIS	≤	10	20
10	<	IIS			30

Note. When ITS = 100% (perfect score), IIP = 30.

GROUP NAME:				TEST DATE:	
TEAM MEMBERS Student ID	Name	Test Score	Pretest	IIS	IIP
TOTAL GROUP IIP					
GROUP AVERAGE IMPROVEMENT POINT (GAIP)					

Figure 5. Group average improvement point (GAIP) worksheet.

Table 26
Criterion of Group Award

Criterion (GAIP)	Group Award
25	Diamond Cup
20	Gold Cup
15	Silver Cup

levels varied from high to average to low and that the average achievement levels of all the groups were approximately the same.

In order for the groups to function at their highest possible level, a survey of "deadly combination" was conducted before the grouping. That is, students had the option to let the instructor know with whom they did not get along well, and it was arranged so that people who had disliked each other did not have to work together for the next 12 weeks.

Fidelity of treatment. To ensure fidelity of treatment, including treatment integrity and treatment differentiation, the researcher used various procedures throughout the duration of the study to ascertain that both the experimental and the control groups follow the protocol they should follow. Treatment integrity refers to "the degree to which a treatment condition is implemented as intended" (Moncher & Prinz, 1991, p. 247), and treatment differentiation refers to "whether treatment conditions differ from one another in the intended manner such that the manipulation of the independent variable actually occurred as planned" (p. 248). The researcher kept frequent face-to-face, email, and telephone communication with the instructor to ensure that the instructor was clear about the necessary steps in each of the instructional methods as well as their differences. In addition, the researcher observed four sessions of class in each group—twice via videotape recording and twice by sitting in the back of the classrooms. The results of the procedures showed that the instructional programs in both the experimental and the control groups were able to be carried out as intended by their individual protocols.

Dependent Variables

There were three main research questions in the study. While all of them had types of instruction as the independent variables, their dependent variables varied from question to question. For Research Question One, the dependent variable was learning motivation; for Research Question Two, it was out-of-class use of learning strategies; for Research Question Three, it was English grammar achievement.

Relationships between Variables

The purpose of this study was to know whether types of instruction, i.e., cooperative learning versus whole-class instruction, influenced outcome (grammar achievement, learning motivation, and use of learning strategies). The relationships between independent variable and dependent variables were causal.

Internal Validity of the Study

In order to be confident that the above mentioned relationships were causal, in other words, in order to be sure that the difference in treatment conditions caused the obtained results, all the threats to internal validity needed to be reasonably controlled. The internal validity of a study is "a judgment that is made concerning the confidence with which plausible rival hypotheses can be ruled out as explanations for the results" (McMillan & Schumacher, 1997, p. 319). According to Creswell (2002) and McMillan and Schumacher (1997), threats to internal validity include history, maturation, statistical regression, selection, mortality, diffusion of treatments, compensatory equalization, compensatory rivalry, pretesting, and instrumentation. The researcher has analytically examined how each of these threats might influence the results and has made efforts to control the threats.

History. This threat concerns unintended or extraneous events occurring between the pretests and the posttests. During the experiment, the researcher had the control group and the experimental group experience the same activities, with the exception of the treatment. For example, when the experimental class went to an English drama show, the control group did as well.

Maturation. People progress and change. They become older, more experienced, and wiser. These changes could have an effect on the outcome of the experiment. However, the subjects of the study were from two freshman classes; the majority of them were between age of 18 and 20 (see Chapter Four for a more detailed description of the sample). Therefore, they would likely mature and develop in a similar way.

Statistical regression. This threat to internal validity can occur if individuals are selected based on extreme scores because the scores tend to gradually regress toward the mean. Two intact college classes were chosen as the experimental and control groups. There were a variety of ability levels in both classes (i.e., a mixture of higher, average and lower achievers). While there were a few students with higher scores in the pretest, the scores were not extreme and did not pose a threat to the validity of the study.

Selection. This refers to individual differences (e.g., intelligence or receptivity to a treatment) that can have an impact on the experimental results. It is reasonable to believe that within each intact class there was a variety of student body in terms of their intelligence, personality and learning styles. In addition, the results of the pretests have shown that there was no significant difference between the two groups in terms of learning motivation, out-of-class use of learning strategies, and English grammar proficiency.

Mortality. When subjects withdraw from the program during the experiment for any reason, it may become difficult to draw conclusions from the data. Fortunately, this was not a concern in the study. The sample size of the study was large enough. Even if a few subjects dropped out, the effects of their withdrawal would be minor.

Diffusion of treatment. When members of the control and experimental groups learn from each other about different treatments, it can create a threat to internal validity. While the researcher recognized the difficulty to completely eliminate the threat, arrangements (e.g., keeping two classes separate in terms of physical proximity, avoiding having interclass activities during the period of the study) were made to keep the two groups as separate as possible.

Compensatory equalization. An inequity occurs if only the experimental group receives a treatment. The inequity could in turn threat the internal validity. To counter the problem, the control group received quality whole-class context-rich instruction for the same duration.

Compensatory rivalry. When variation in treatments is openly pronounced, compensatory rivalry could occur between the experimental and the control groups. The researcher made efforts to avoid the threat by attempting to reduce the awareness and expectations of the presumed benefits of the experimental treatment.

Pretesting. Another potential threat to the validity of the study was that subjects could remember their responses from the pretest and thus affected the outcome of the posttest. In terms of the achievement measurement, the pretest and the posttest used different forms, so there was no concern at all for the threat. In terms of the questionnaire, the posttest did not take place until 12 weeks after the pretest. Many instructions and

learning took place in this period, so there was little chance the subjects remembered responses from the pretest.

Instrumentation. The threat refers to inadequate demonstration of the reliability and validity of measurement tools. There was no concern for this threat because all the instruments in this study had reasonable reliability and validity.

Data Collection

The subjects were 84 students from two intact college EFL classes in Taiwan. One of the classes received cooperative learning in the form of the STAD while the other received whole-class instruction. The experimental group and the control group were measured twice: before and after the 12-week treatment. The pretests included an English grammar proficiency test and a motivational learning strategies questionnaire; the posttests included a grammar achievement test and the same questionnaire.

Data Analyses

Descriptive Analyses

Means and frequency distributions were calculated to determine the subjects' general performance in the grammar tests as well as their motivation and use of learning strategies. For posttest scores of the above-mentioned variables, adjusted means were also obtained from the analyses of covariance (ANCOVAs) to present a more comprehensive picture of student attitude, behaviors, and performance.

Inferential Analyses

With regard to motivation and use of learning strategies, a one-way multivariate analysis of covariance (MANCOVA) on the posttest scores with pretest scores as the covariates was conducted to see if there was a significant difference between the

experimental group and the control group on each set of the dependent variables. Potential pretreatment differences between groups were controlled with the use of covariates. The covariates, which were systematically related to the dependent variables, were used to reduce the estimate of random or error variance in the dependent measures, to eliminate systematic bias by adjusting group means on the dependent variables, and to increase the power of the statistical tests (Huck, 2004). When a MANCOVA was found significant, investigation was followed up with a univariate ANCOVA on each of the dependent variables to see if there was a significant difference between the experimental group and the control group.

With regard to grammar achievement, a univariate ANCOVA was used to determine if a significant difference existed between the two comparison groups.

The general statistical hypothesis for each analysis was as follows:

$H_0: \mu_1' = \mu_2'$

$H_1: \mu_1' \neq \mu_2'$

An alpha level of $\alpha = .05$ was selected. Because the researcher was not necessarily concerned about Type I error, an alpha level of $\alpha = .05$ was chosen by convention. The desired level of power was estimated to be .80 using the formula 1-4(α). The researcher had assumed a large effect size of 1.00, a power level of .80, an alpha level of .05, and treatment level of 2. According to Hinkle, Wiersma, and Jurs (1994, p. 634), the sample size needed is 17 per group. The exact sample size of 42 per group in the study has exceeded this requirement.

Tests for Assumptions Underlying ANCOVA

The following are the main assumptions underlying ANCOVA. Care was taken in this study to ensure all the assumptions were satisfied. The assumptions are discussed below.

1. The observations are normally distributed on the dependent variable in each group. This is the normality assumption. Violation of this assumption will only minimally affect the sampling distribution of F (Stevens, 1999) and "has negligible consequences on Type-I and Type-II error probabilities" (Glass & Hopkins, 1996, p. 403). That means the procedure of ANCOVA is robust with regard to the violation of the normality assumption. The assumption was therefore not tested.

2. The population variances across groups are equal. This is the homogeneity of variance assumption. If the sample group sizes are equal or approximately equal, i.e., the sample size of the largest group is no more than 1.5 times the sample size of the smallest group, the procedure of ANCOVA is robust to the violation of this assumption (Glass & Hopkins, 1996; Stevens, 1999). The sizes of the groups were equal in this study.

3. The observations within groups are not influenced by each other. This is the independence assumption. But as Stevens (1999) and Glass and Hopkins (1996) indicated, in teaching methods studies, especially those involved discussion among group members, dependence among subjects is inevitable. In such a situation, Stevens (1999) suggested using the group means as the unit of analysis. This study used the group means, rather than individual scores, as the units of analysis.

4. The dependent variable and the covariate are related in a linear fashion. This is the linearity assumption. In the present study, the assumption was checked with

scatterplots, residual plots, and the Tolerance and VIF values to see whether there were likely problems with multicollinearity (Leech, Barrett, & Morgan, 2005; Lomax, 2001b).

5. The slopes of the regression lines for the covariate in relation to the dependent variable must be the same for each group, i.e., the regression lines should be parallel. This is the assumption of homogeneity of regression, or the assumption of homogeneity of regression slopes. The assumption is one of the most important ANCOVA assumptions. In this study, F tests on the interaction of the independent variables with the covariates were conducted to ascertain whether the assumption was satisfied (Leech, Barrett, & Morgan, 2005).

6. Another important assumption for ANCOVA is that the covariate is measured without error, the violation of which can cause the treatment effects to be seriously biased (Lomax, 2001a). To prevent possible violation, Lomax (2001b) and Pedhazur (1997) suggested constructing measures of the covariates that have good reliabilities prior to the study or computing the reliabilities of the covariates from previous research. Lomax (2001b) further suggested considering the validity of the covariates. In the present study, the researcher has made efforts to locate and construct instruments that had sound validity and reliability (see earlier discussions in the Validity and Reliability sections) and thus minimized the possibility of violating this assumption.

Tests for Assumptions Underlying MANCOVA

The following are the main assumptions underlying MANCOVA. Efforts were made to ensure all the assumptions were met. The assumptions are discussed as follows.

1. The observations on the dependent variables follow a multivariate normal distribution in each group. This is the multivariate normality assumption (Stevens, 1996).

MANCOVA is robust to its violation if the sample group sizes are nearly equal, i.e., largest/smallest < 1.5 (Leech, Barrett, & Morgan, 2005). The group sizes in the present study were the same. The assumption was thus not tested.

2. The population variances across groups are equal. This is the homogeneity of variance assumption. It was not tested as MANCOVA is robust to its violation if the group sizes are approximately equal, i.e., largest/smallest < 1.5 (Leech, Barrett, & Morgan, 2005).

3. The observations are independent. See the independence assumption for ANCOVA.

4. Covariances between pairs of dependent variables are approximately equal for all groups. This is the homogeneity of covariance matrices assumption. Leech, Barrett, and Morgan (2005) suggested using the Box's test of equality of covariance matrices to check the assumption. As a result of whether the assumption is met, different multivariate tests should be used for MANCOVA. For example, Wilks' Lambda is most commonly used when the assumption is satisfied. Pillai's Trace is the best option when the assumption is violated and the group sizes are similar. In this study multivariate statistics for MANCOVA were chosen carefully based on these guidelines.

Exploratory Analyses

Factorial ANCOVAs were conducted to see if there were significant interactions between the effects of prior English level and instruction on each of the dependent variables. A significant interaction was followed up with simple effect analyses in order to examine what the cause of the interaction was. Interaction plots were used to help interpret the results.

In addition, ANCOVAs on the grammar posttest scores of various cognitive levels were conducted to examine the effects of cooperative learning. Pearson product moment correlations were used to investigate the relationships among prior English ability level, gender, motivation, use of learning strategy, and grammar achievement.

Chapter Summary

This chapter contains a description of the subjects, the instrumentation, and the procedures used for the study. The subjects are discussed in terms of the sampling procedure, sample size, sample characteristics, and external validity. The instrumentation is discussed in terms of the instruments' purposes, technical characteristics, scoring procedures, as well as score interpretation. The procedures are discussed in terms of the research design, the data collection methods, and the data analysis methods.

The study used a quasi-experimental pretest-posttest comparison group research design. Eighty-four subjects from two intact college EFL classes in Taiwan were involved in a 12-week experiment. One class received English grammar instruction through cooperative learning and the other class through whole-class teaching. The data were collected once at pretest and once at posttest. The instruments employed to measure students' motivation, out-of-class learning strategy utilization, and grammar achievement included the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993), a grammar pretest derived from the GEPT test bank, as well as a grammar posttest developed by the researcher of the present study.

To determine the subjects' general performance in the grammar tests as well as their motivation and learning strategy utilization beyond the class context, means and frequency distributions were calculated. For posttest scores of these variables, adjusted

means were also obtained from the analyses of covariance (ANCOVAs) to show a more comprehensive depiction of student motivation, learning behaviors, and cognitive performance.

A univariate ANCOVA was conducted to examine if there was a significant difference between the experimental group and the control group on grammar achievement. In terms of motivation and learning strategy use, one-way multivariate analyses of covariance (MANCOVAs) were conducted to see if any significant difference existed between the two comparison groups. When a MANCOVA was found significant, the investigation was followed up with a univariate ANCOVA on each of the dependent variables (i.e., self-efficacy, task value, use of elaboration skills, and use of peer collaboration) to determine if there was a significant difference between the two groups.

In order to investigate whether there were significant interaction effects between type of instruction and prior English ability level on grammar achievement, self-efficacy, task value, elaboration, and peer collaboration, factorial ANCOVAs were conducted. A significant interaction was followed up with simple effect analyses to examine what the cause of the interaction was. Interaction plots were employed to facilitate the interpretation of the results.

In addition, ANCOVAs were performed to understand the differential effects of cooperative learning at different cognitive levels. Pearson product moment correlations were used to examine the relations among prior English ability level, gender, grammar achievement, motivation in terms of task value and self-efficacy, and use of learning strategies in terms of elaboration and peer learning. The results of these data analyses will be presented in Chapter Four.

CHAPTER FOUR

RESULTS

A review of literature (see Chapter Two) showed some efforts had been made in Taiwan to enhance the effectiveness of EFL instruction through cooperative learning. Yet it appeared that more systematic studies on the topic were needed before conclusions could be made regarding the effects of cooperative learning on Taiwanese learners of English. Hence, the purpose of this study was to examine the differential effects of the cooperative learning and whole-class instruction methods on motivational and cognitive measures. Three major research questions were asked. (1) How does motivation differ between the group receiving cooperative learning and the group receiving whole-class instruction? (2) How does utilization of learning strategies beyond class settings differ between the group receiving cooperative learning and the group receiving whole-class instruction? (3) How does grammar achievement differ between the group receiving cooperative learning and the group receiving whole-class instruction? Based on the results to the above questions, the following exploratory questions were asked. They are listed as Exploratory Questions A to E. (A) Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on motivation? If so, what is the cause of the interaction? (B) Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on out-of-class utilization of learning strategies? If so, what is the cause of the interaction? (C) Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on grammar achievement? If so, what is the

cause of the interaction? (D) How does student achievement differ between the cooperative learning group and the whole-class instruction group in terms of different cognitive levels? (E) What are the relationships among prior English level, gender, grammar achievement, task value, self-efficacy, use of elaboration strategies, and out-of-class peer collaboration behaviors? The findings of this study will attempt to answer the questions and present information that will be of assistance to EFL instructors in Taiwan when they come to selections of EFL practice.

Data were collected from 84 students (42 from each group) via the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) in order to better understand student motivation and use of learning strategies. Data were also collected by means of scores from an English grammar pretest and a grammar posttest (see Chapter Three for a detailed description of these tests) to examine the effects of cooperative learning on student achievement. This chapter represents the statistical analyses of the data.

The chapter is organized into the following sections: descriptive analyses, inferential analyses, exploratory analyses, and chapter summary.

Descriptive Analyses

This section presents descriptive statistical information on the sample, the MSLQ questionnaires, and the grammar tests. The descriptive results of the MSLQ questionnaires are further discussed in terms of the self-efficacy for learning and performance scale and the task value scale, which were used to measure student motivation, and the elaboration scale and the peer learning scale, which were employed

to measure students' use of learning strategies beyond classroom settings. The grammar tests are discussed in terms of the pretest and the posttest.

The Sample

A sample of 84 students was distributed evenly among the two intact college classes, with 42 subjects in each group. These two classes were randomly assigned treatment conditions. One group received cooperative learning and the other whole-class instruction. Table 27 presents the demographic information of the total sample, the cooperative learning group and the whole-class instruction group.

Table 27
Demographic Information of the Sample

	Total Sample		Cooperative Group		Whole-Class Group	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Gender						
Male	22	26.2	10	23.8	12	28.6
Female	62	73.8	32	76.2	30	71.4
Total	84	100.0	42	100.0	42	100.0
Age						
18	48	57.1	24	57.1	24	57.1
19	11	13.1	6	14.3	5	11.9
20	16	19.0	7	16.7	9	21.4
21	3	3.6	2	4.8	1	2.4
22	6	7.2	3	7.1	3	7.2
Total	84	100.0	42	100.0	42	100.0

The total sample was comprised of approximately one fourth (22) male and three fourths (62) female. The total sample was quite homogeneous in age. All subjects were between the ages of 18 and 22. Around 90% (75) of the sample were between the ages of 18 and 20.

Regarding the subjects' gender, both groups, like the total sample, had approximately one fourth male (10 in the cooperative learning group and 12 in the whole-class instruction group) and three fourths female (32 and 30 respectively). Regarding the subjects' age, the cooperative group (M = 18.90, SD = 1.27) and the whole-class group (M = 18.90, SD = 1.25) also resembled the total sample (M = 18.90, SD = 1.25). Both groups had approximately 90% of subjects between the ages of 18 and 20.

With regard to demographics, the data indicated that the two groups were similar in terms of gender distribution and age.

MSLQ Questionnaires

Four scales from the MSLQ questionnaire were used to measure the sample's motivation and out-of-class use of learning strategies. Motivation was measured in terms of self-efficacy and task value. Use of learning strategies was measured in terms of elaboration and peer learning. Possible scale scores ranged from 1 to 7. The descriptive statistical information is presented in Table 28. The discussions followed hereafter will use the criteria set in Table 17 for score interpretation.

Motivation Scales

Self-efficacy pretest. The self-efficacy for learning and performance scale measured the subjects' expectations and their confidence levels in carrying out tasks. As presented in Table 28, the total sample on average showed moderate self-efficacy (M = 3.80) in the pretest. Variation among the total sample was moderate with a standard

Table 28
Mean Scores for the MSLQ Scales

Group	n	Pretest		Posttest		
				Obtained		Adjusted
		Mean	SD	Mean	SD	Mean
Self-Efficacy for Learning and Performance						
Cooperative	42	3.69	1.39	4.79	1.04	4.87
Whole-Class	42	3.91	1.43	4.02	1.34	3.94
Total Sample	84	3.80	1.40	4.41	1.25	4.41
Task Value						
Cooperative	42	3.57	1.09	5.16	.82	5.26
Whole-Class	42	3.81	1.16	3.68	1.12	3.59
Total Sample	84	3.69	1.13	4.42	1.23	4.43
Elaboration						
Cooperative	42	3.73	.88	4.96	.92	4.97
Whole-Class	42	3.75	.91	3.81	.89	3.80
Total Sample	84	3.74	.89	4.39	1.07	4.39
Peer Learning						
Cooperative	42	2.51	.52	4.02	.73	4.07
Whole-Class	42	2.71	.43	2.74	.47	2.68
Total Sample	84	2.61	.49	3.38	.89	3.38

deviation of 23% of the scoring range. The attitude of the cooperative learning group and the whole-class instruction group appeared somewhat similar, both showing moderate self-efficacy (M = 3.69 and 3.91 respectively). Like the total sample, variations within each group were also moderate.

Self-efficacy posttest. On the self-efficacy posttest, the total sample on average scored slightly higher than in the pretest. The mean score (M = 4.41) still indicated moderate self-efficacy. Variation among the total sample was moderate with a standard deviation of 21% of the scoring range. The cooperative learning group showed enhanced self-efficacy in the posttest. After the treatment, the level of self-efficacy raised from moderate to high (M = 4.79). On the other hand, that of the whole-class instruction group

remained moderate (M = 4.02). Variation within the cooperative group was small to moderate with a standard deviation of 17% of the scoring range and that within the whole-class group was moderate with a standard deviation of 22% of the scoring range.

Task value pretest. The task value scale measured how the subjects perceived the course activities and materials in terms of their interest, importance, and relevance. As indicated in Table 28, the total sample on average showed moderate task value (M = 3.69) in the pretest. Variation among the total sample was moderate with a standard deviation of about 20% of the scoring range. The task value of the cooperative learning group and the whole-class instruction group appeared somewhat similar, both showing moderate task value (M = 3.57 and 3.81 respectively). Variations within each group were also moderate.

Task value posttest. On the task value posttest, the total sample on average scored slightly higher than in the pretest. The mean score (M = 4.42) indicated a moderate level of task value. Variation among the total sample was moderate with a standard deviation of 21% of the scoring range. The cooperative learning group showed enhanced task value on the posttest. After the treatment, the level of task value increased from moderate to high. Variation within the cooperative group was small to moderate with a standard deviation of 14% of the scoring range. On the other hand, the whole-class instruction group's task value remained moderate. Variation within the whole-class group was moderate with a standard deviation of about 20% of the scoring range.

Use of Learning Strategies Scales

Elaboration pretest. The elaboration scale measured the subjects' utilization of learning strategies, such as paraphrasing, summarizing, and synthesizing, beyond the

class context. In the elaboration pretest, the total sample on average showed moderate use of elaboration strategies (M = 3.74). Variation among the total sample was small to moderate with a standard deviation of 15% of the scoring range. The cooperative learning group and the whole-class instruction group were almost identical, both showing moderate use of elaboration (M = 3.73 and 3.75 respectively). Variations within each group were small to moderate.

Elaboration posttest. On the elaboration posttest, the total sample on average scored slightly higher than in the pretest. The mean score (M = 4.39) still indicated only moderate use of elaboration. Variation among the total sample was moderate with a standard deviation of about 20% of the scoring range. The cooperative learning group showed more use of elaboration in the posttest. After the treatment, the level of elaboration strategy use raised from moderate to frequent (M = 4.96). On the other hand, the whole-class instruction group's use of elaboration showed little change and remained moderate (M = 3.81). Variations within each group were small to moderate with a standard deviation of 15% of the scoring range.

Peer learning pretest. The peer learning scale measured the subjects' use of peer collaboration strategies in learning beyond the class context. The total sample on average showed infrequent use of peer learning strategies (M = 2.61) in the pretest. Variation among the total sample was small with a standard deviation of 8% of the scoring range. The cooperative learning group and the whole-class instruction group were similar in this regard, both showing infrequent use of peer learning (M = 2.51 and 2.71 respectively) with small variations within each group.

Peer learning posttest. The total sample on average showed infrequent use of peer learning (M = 3.38). Variation among the total sample increased to a small to moderate level with a standard deviation of 15% of the scoring range. The cooperative learning group showed more use of peer learning in the posttest. After the treatment, the peer learning strategy use increased from infrequent to moderate (M = 4.02). Variation within the group was relatively small with a standard deviation of 12% of the scoring range. On the other hand, the whole-class instruction group's use of peer learning strategies showed little change and remained infrequent (M = 2.74). Variation within the group was small with a standard deviation of 8% of the scoring range.

Summary on the MSLQ Descriptive Analyses

In this study, motivation was measured in terms of self-efficacy and task value. Before the treatment, the cooperative group and the whole-class group all showed moderate levels of self-efficacy and task value. The subjects demonstrated only a moderate level of self-efficacy in completing course tasks. While they had some expectation and confidence in carrying out their learning tasks, the students' level of anticipation and self-belief was not high. In addition, the subjects on average attached only moderate levels of value to the course task. While they perceived some value upon the course activities and materials, they did not think highly in terms of their interest, importance, and relevance.

After the treatment, the whole-class group's self-efficacy and task value showed little change and remained only at the moderate level. In contrast, the cooperative learning group showed improvement on both self-efficacy and task value, which were enhanced from moderate levels on the pretest to high levels on the posttest. In other

words, after the treatment, the cooperative group on average demonstrated higher expectation and self-confidence in learning and performing and perceived more highly of the course task.

Use of learning strategies was measured in terms of elaboration and peer learning. Before the treatment, the cooperative group and the whole-class group all showed moderate use of elaboration strategies and infrequent use of peer learning strategies. After the treatment, the whole-class group continued to show moderate use of elaboration strategies and infrequent use of peer learning strategies. The cooperative learning group, on the other hand, demonstrated increased use of elaboration and peer learning. The use of elaboration strategies was increased from a moderate to a frequent level, and the use of peer learning strategies from an infrequent to a moderate level.

Grammar Tests

The sample received a grammar pretest and a grammar posttest before and after they received their treatment (i.e., cooperative learning or whole-class instruction). Table 29 shows the mean scores of the test results for each group and the total sample.

Table 29
Mean Scores for the Grammar Tests

Group	n	Pretest		Posttest		
				Obtained		Adjusted
		Mean	SD	Mean	SD	Mean
Cooperative	42	36.38	15.35	69.00	18.51	70.96
Whole-Class	42	40.48	20.44	66.62	20.20	64.67
Total Sample	84	38.43	18.08	67.81	19.29	67.82

Pretest

Possible pretest scores ranged from 0 to 100. The total sample on average answered slightly more than one third of the items correctly. Based on the General English Proficiency Test (GEPT) standard, i.e., 67.00 as the passing benchmark, this would indicate a very unsatisfactory failing score. Variation among the total sample was large, with scores ranging about 80 points. The performances of the cooperative learning group and the whole-class instruction group appeared somewhat similar (36.38 vs. 40.48), both showing performances far below the benchmark. Variations within each group were large, with scores ranging across about 65 points in the cooperative group and 81 points in the whole-class group.

Posttest

Possible posttest scores ranged from 0 to 100. The total sample on average answered two thirds of the items correctly. This was similar to the benchmark of 67.00. Variation among the total sample was large, with scores ranging about 73 points. The performances of the cooperative learning group and the whole-class instruction group appeared somewhat similar (69.00 vs. 66.62), showing performance slightly above or just below the benchmark. Variations within each group were still large, with scores ranging across about 59 points in the cooperative group and 70 points in the whole-class group.

Inferential Analyses

One-way multivariate analyses of covariance (MANCOVAs) and one-way univariate analyses of covariance (ANCOVAs) were conducted to answer Research Questions One to Three. An alpha level of .05 was used for all statistical tests. Covariates

were used to reduce the error variances on the dependent variables and to increase statistical power (Huck, 2004).

Research Question One

The first research question was: How does motivation differ between the group receiving cooperative learning and the group receiving whole-class instruction? To answer this question, first a MANCOVA was conducted to examine the use of multiple dependent variables. All of the assumptions underlying the use of MANCOVA were satisfied. The assumption of multivariate normality and the assumption of homogeneity of variance were not tested as MANCOVA is robust to the violations if groups are of nearly equal size (Leech, Barrett, & Morgan, 2005). The assumption of independence of observations was assumed to be true. The assumption of homogeneity of covariance matrices was tested with the Box's test of equality of covariance matrices. No significance difference was found ($F_{3, 1210320} = 2.38$, $p = .068$) and the assumption was met. According to Leech, Barrett, and Morgan (2005), under most conditions when the assumptions are satisfied, Wilks' Lambda is an appropriate multivariate statistic to use for MANCOVA. Results found a significant difference between the cooperative learning and the whole-class groups (Wilks' Lambda = .23, $F_{2, 79} = 135.73$, $p = .000$). The effect size[1] was large (eta squared = .78) and the observed power[2] was excellent at 1.00. Hence subsequent univariate ANCOVAs, one on self-efficacy for learning and performance and the other on task value, were performed.

[1] An eta squared of .01, .06, and .14 denotes small, medium, and large effect sizes, respectively (Cohen, 1988; Green, Salkind, & Akey, 2000, p. 190).
[2] A power of .70 is generally considered adequate and a power of .90 excellent (Stevens, 1999, p. 123).

Self-Efficacy

An examination of the results in Table 30 indicated that the cooperative group's average posttest score on self-efficacy scale (adjusted mean = 4.87) was significantly higher ($F_{1, 81} = 58.77$, p = .000) than that of the whole-class group (adjusted mean = 3.94). (For a detailed presentation of the mean scores and adjusted mean scores of the self-efficacy, task value, elaboration, and peer learning scales, see Table 28.) The observed power was excellent at 1.00, and the effect size was large at .42.

Table 30
One-Way ANCOVA on Self-Efficacy Posttest Scores with Self-Efficacy Pretest Scores as Covariate

Source	SS	df	MS	F	p
Pretest	87.21	1	87.21		
Group	18.16	1	18.16	58.77	.000
Error	25.03	81	.31		
Total	130.40	83	`		

All of the assumptions underlying the use of ANCOVA were satisfied. The assumption of normality was not tested as the procedure is robust to its violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82} = .06$, p = .815, indicating equivalency of variances across the groups. The scatterplot and the residual plot were used to test the linearity. The Tolerance and VIF values equaled 1. A linear relationship between the dependent variable (i.e., the posttest) and the covariance was thus found and that the assumption of linearity met. In

addition, the assumption of homogeneity of regression was tested. The observed value was $F_{14, 32} = .82$, $p = .641$. The result showed that the lines of the dependent variable that had been regressed on the covariate within each group were parallel and the assumption of homogeneity of regression was met.

Task Value

An examination of the results in Table 31 indicated that the cooperative group's posttest scores on task value scale (adjusted mean = 5.26) were significantly higher ($F_{1, 81} = 221.40$, $p = .000$) than those of the whole-class group (adjusted mean = 3.59). The observed power was excellent at 1.00, and the effect size was large (eta squared = .73).

Table 31
One-Way ANCOVA on Task Value Posttest Scores with Task Value Pretest Scores as Covariate

Source	SS	df	MS	F	p
Pretest	47.11	1	47.11		
Group	57.58	1	57.58	221.40	.000
Error	21.07	81	.26		
Total	125.75	83	`		

All the assumptions pertaining to the use of ANCOVA were satisfied. The assumption of normality was not tested as the procedure is robust to its violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82} = .01$, $p = .974$, indicating equivalency of variances across the groups.

The scatterplot and the residual plot were used to test the linearity. The Tolerance and VIF values equaled 1. It was thus concluded that the assumption of linearity was not violated. In addition, the assumption of homogeneity of regression was tested. The observed value was $F_{11, 48} = 1.95$, $p = .056$. The result showed that the regression lines of the dependent variable that had been regressed on the covariate within each group were parallel and the assumption of homogeneity of regression was met.

Research Question Two

The second research question was: How does utilization of learning strategies beyond class settings differ between the group receiving cooperative learning and the group receiving whole-class instruction? To answer this question, first a MANCOVA was conducted to examine the use of multiple dependent variables. All of the assumptions underlying the use of MANCOVA were satisfied. The assumption of multivariate normality and the assumption of homogeneity of variance were not tested as MANCOVA is robust to the violations when group sizes are approximately equal (Leech, Barrett, & Morgan, 2005). The assumption of independence of observations was assumed to be true. The assumption of homogeneity of covariance matrices was tested with the Box's test of equality of covariance matrices. Significant difference was found ($F_{3, 1210320} = 3.13$, $p = .025$) and the assumption thus violated. According to Leech, Barrett, and Morgan (2005), if the assumption is violated but the group sizes are similar (which was the case of the present study), MANCOVA would still be robust if Phillai's Trace is used as the multivariate statistic. Phillai's Trace, therefore, was used for this MANCOVA. Results indicated a significant difference between the cooperative learning and the whole-class groups (Phillai's Trace = .69, $F_{2, 79} = 88.96$, $p = .000$). The effect size was large (eta

squared = .69) and the observed power was excellent at 1.00. As a result, subsequent univariate ANCOVAs, including one on elaboration and the other on peer learning, were performed.

Elaboration

An examination of the results in Table 32 indicated a statistically significant ($F_{1, 81}$ = 115.54, p = .000) treatment effect for the use of elaboration strategies, with the learners in the cooperative group on average (adjusted mean = 4.97) utilizing more elaboration strategies than the learners in the whole-class group (adjusted mean = 3.80). The observed power was excellent at 1.00, and the effect size was large (eta squared = .59).

Table 32
One-Way ANCOVA on Elaboration Posttest Scores with Elaboration Pretest Scores as Covariate

Source	SS	df	MS	F	p
Pretest	45.71	1	45.71		
Group	28.98	1	28.98	115.54	.000
Error	20.32	81	.25		
Total	95.01	83			

All the assumptions underlying the use of ANCOVA were met. The assumption of normality was not tested as the procedure is robust to its violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82}$ = .20, p = .653, indicating equivalency of variances across the groups. The linearity between the dependent variable (i.e., the posttest) and the covariance was tested with the

scatterplot, the residual plot, as well as the Tolerance and VIF values (both equaled 1). The results indicated that the assumption of linearity was satisfied. In addition, the assumption of homogeneity of regression was tested. The observed value was $F_{15, 47}$ = 1.50, p = .146. The result showed that the lines of the dependent variable that had been regressed on the covariate within each group were parallel and the assumption of homogeneity of regression was met.

Peer Learning

An examination of the results in Table 33 indicated that the average cooperative group's posttest score on the peer learning scale (adjusted mean = 4.07) was significantly higher ($F_{1, 81}$ = 124.04, p = .000) than that of the whole-class group (adjusted mean = 2.68). The observed power was excellent at 1.00, and the effect size was large (eta squared = .61).

Table 33
One-Way ANCOVA on Pear Learning Posttest Scores with Pear Learning Pretest Scores as Covariate

Source	SS	df	MS	F	p
Pretest	1.29	1	1.29		
Group	38.62	1	38.62	124.04	.000
Error	25.22	81	.31		
Total	65.06	83	`		

None of the assumptions pertaining to the use of ANCOVA in this situation was violated. The assumption of normality was not tested as the procedure is robust to its violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes

were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82} = 3.37$, $p = .070$, indicating equivalency of variances across the groups. The scatterplot and the residual plot were used to test the linearity. The Tolerance and VIF values equaled 1. It was thus concluded that there was a linear relationship between the dependent variable (i.e., the posttest) and the covariance and that the assumption of linearity was satisfied. Additionally, the assumption of homogeneity of regression was tested. The observed value was $F_{5, 70} = 1.31$, $p = .269$. The result showed that the lines of the dependent variable that had been regressed on the covariate within each group were parallel and the assumption of homogeneity of regression was met.

Research Question Three

The third research question was: How does grammar achievement differ between the group receiving cooperative learning and the group receiving whole-class instruction? This question was answered with a one-way ANCOVA on the grammar posttest scores using the grammar pretest scores as the covariate. The results from the one-way ANCOVA are presented in Table 34.

Table 34
One-Way ANCOVA on Grammar Posttest Scores with Grammar Pretest Scores as Covariate

Source	SS	df	MS	F	p
Pretest	23711.86	1	23711.86		
Group	820.25	1	820.25	10.44	.002
Error	6363.75	81	78.57		
Total	30895.85	83			

An examination of the results indicated a statistically significant ($F_{1, 81} = 10.44$, $p = .002$) treatment effect on grammar achievement, with the learners in the cooperative group obtaining a higher average grammar posttest score (adjusted mean = 70.96) than the learners in the whole-class group (adjusted mean = 64.67). (See Table 29 for a detailed display of the grammar mean scores and the adjusted mean scores.) The observed power was high at .89, and the effect size was medium-to-large (eta squared = .11).

None of the assumptions underlying the use of ANCOVA was violated. The assumption of normality was not tested as the procedure is robust to its violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82} = 3.28$, $p = .074$, indicating equivalency of variances across the groups. The scatterplot and the residual plot were used to test the linearity. The Tolerance and VIF values equaled 1. It was thus concluded that the assumption of linearity was met. In addition, the assumption of homogeneity of regression was tested. The observed value was $F_{14, 50} = 1.42$, $p = .181$. The result showed that the lines of the dependent variable that had been regressed on the covariate within each group were parallel, and thus the assumption of homogeneity of regression was satisfied.

Summary on the Inferential Analyses

Research Questions One to Three investigated the differential effects of two instructional approaches, cooperative learning and whole-class instruction, on motivation, learning strategies, and academic achievement. Areas of exploration included

self-efficacy on learning and performance, task value, use of elaboration strategies, use of peer collaboration strategies, and grammar achievement. The results of all these research questions were found to be significant in favor of cooperative learning as more effective than whole-class instruction.

Exploratory Analyses

As a result of the inferential analyses, Exploratory Questions A to E were investigated. Factorial ANCOVAs, simple effect analyses, one-way ANCOVAs, and Pearson product moment correlations were conducted to answer these questions.

In order to answer Exploratory Questions A, B, and C, two approaches were taken to divide the subjects into different English ability levels based on their pretest performance. The first approach used the GEPT score interpretation as a guideline. The GEPT test adopted a pass/fail system. In a GEPT reading test, a test taker needed to answer 67% of items correctly to pass the test. The researcher hence operationally defined those who scored 67/100 and above in the pretest as higher-proficiency students. There were 9 subjects in this category, 3 from the cooperative learning group and 6 from the whole-class group. When a subject was not even half way to the GEPT passing score of 67 (i.e., he or she earned a 33 or less on the pretest), he or she was categorized as a lower-proficiency student. They totaled 42, with 21 in each group. Those who scored lower than 67 and higher than 33 were categorized as medium-proficiency students. There were 33 of them, 18 in the cooperative learning group and 15 in the whole-class group.

The second approach used the subjects' relative ranking in class to categorize them into various levels. As previously described, the subjects in the cooperative learning

class were grouped into 10 cooperative teams, with 4 students in most teams and 5 in 2 teams. Based on the rationale that students' learning motivation and effort could be affected by how they perceived their relative standing among classmates and among teammates (Bandura, 1993; Covington, 1992), and also based on the rationale that in each cooperative learning teams the student with a higher pretest score would tend to have more opportunity to elaborate and explain the concept of English grammar to the group members (Slavin, 1995), the researcher decided that it was worth exploring to categorize the students into quartiles and identify them as higher (the top quartile), medium (the middle two quartiles), and lower (the bottom quartile) ranking students for the purpose of analyses on Exploratory Questions A, B, and C.

The intent of conducting Exploratory Questions A, B, and C was to examine if there was a significant interaction effect between English level and instruction type, and, if so, what the cause of the interaction was. Main effects of English level and instruction were not the intended areas of exploration. The following data analyses and interpretation of results, thus, focused on interaction and subsequent simple effects only.

Exploratory Question A

The first exploratory question was: Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on motivation? If so, what is the cause of the interaction? Motivation was measured in terms of self-efficacy for learning and performance and task value. Each scale was analyzed with two, two-way ANCOVAs with instruction (CL vs. WC) and ability level (higher, medium, lower) as the independent variables. The first factorial ANCOVA used the GEPT standard to define

ability levels; the second used students' relative ranking in class as the operational definition.

Self-Efficacy

Results from the two factorial ANCOVAs on self-efficacy (see Table 35) indicated no significant interaction between prior GEPT level and type of instruction ($F_{2, 77} = 2.08$, $p = .132$) while the main effects of prior GEPT level and instruction were both significant ($F_{2, 77} = 5.06$, $p = .009$ and $F_{1, 77} = 44.21$, $p = .000$ respectively). The results revealed that the effect of instruction on self-efficacy did not depend on which GEPT level was being considered, and vice versa.

Table 35
Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Self-Efficacy

Source	SS	df	MS	F	p
Ability Defined by GEPT Standard					
Pretest	21.18	1	21.18		
GEPT level	2.76	2	1.38	5.06	.009
Instruction	12.07	1	12.07	44.21	.000
GEPT level x Instruction	.14	2	.57	2.08	.132
Error	21.02	77	.27		
Ability Defined by Relative Ranking in Class					
Pretest	6.39	1	6.39		
Ranking	4.31	2	2.16	12.24	.000
Instruction	19.87	1	19.87	112.82	.000
Ranking x Instruction	6.96	2	3.48	19.76	.000
Error	13.56	77	.18		

On the other hand, there was a significant interaction between the effects of prior class ranking and type of instruction on self-efficacy ($F_{2, 77} = 19.76$, $p = .000$). The

observed power was excellent at 1.00, and the effect size was large (eta squared = .34). The statistically significant interaction indicated that the effect of instruction on self-efficacy depended on which prior ranking level was being considered.

To examine where the significant ranking by instruction interaction effect on self-efficacy took place, three simple effect analyses were conducted (see Levine, 1991). The analyses included one on the two group means within the higher ranking, another on the two group means within the medium ranking, and the other on the two group means within the lower ranking. The simple effect analyses revealed that type of instruction influenced the self-efficacy of the higher ranking students ($F_{1, 77} = 26.36$, $p = .000$), medium ranking students ($F_{1, 77} = 9.53$, $p = .003$), and lower ranking students ($F_{1, 77} = 92.19$, $p = .000$). However, results from the interaction plot (see Figure 6) indicated that even though significant differences existed between the two instruction groups across all three ranking levels, the difference between the middle groups appeared to be relatively smaller than the differences of the other two groups.

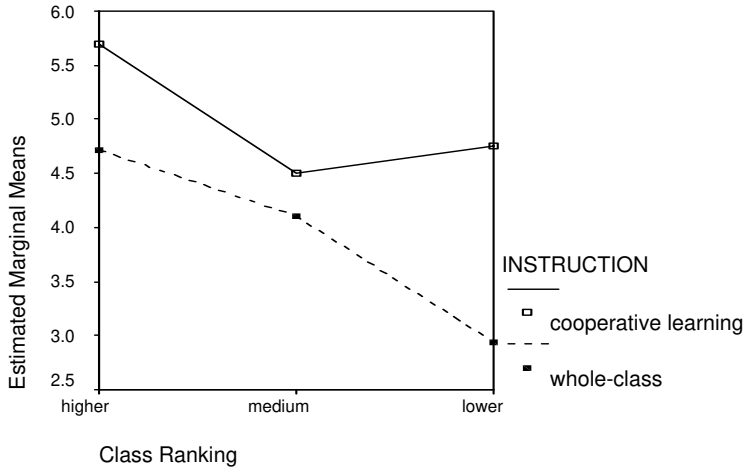

Figure 6. Interaction plot for ranking and instruction on self-efficacy.

Task Value

Results from the two, two-way ANCOVAs on task value appear in Table 36. An examination of the results indicated no significant interaction between GEPT level and instruction ($F_{2, 77} = 1.07$, $p = .348$). Significant main effects were found for instruction ($F_{1, 77} = 137.77$, $p = .000$) but not for GEPT level ($F_{2, 77} = 1.91$, $p = .156$). The results revealed that the effect of instruction on task value did not depend on which GEPT level was being considered, and vice versa.

On the other hand, there was a significant interaction between the effects of prior class ranking and instruction ($F_{2, 77} = 3.85$, $p = .026$) on task value. The observed power was reasonable (.68), and the effect size was medium-to-large (eta squared = .09). The statistical significant interaction indicated that the effect of instruction on task value depended on which prior class ranking level was being considered.

Table 36
Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Task Value

Source	SS	df	MS	F	p
Ability Defined by GEPT Standard					
Pretest	31.59	1	31.59		
GEPT level	.96	2	.48	1.91	.156
Instruction	34.69	1	34.69	137.77	.000
GEPT level x Instruction	.54	2	.27	1.07	.348
Error	19.39	77	.25		
Ability Defined by Relative Ranking in Class					
Pretest	4.83	1	4.83		
Ranking	5.13	2	2.56	13.52	.000
Instruction	47.98	1	47.98	252.90	.000
Ranking x Instruction	1.46	2	.73	3.85	.026
Error	14.61	77	.19		

To examine where the significant ranking by instruction interaction effect took place, three simple effect analyses were conducted. They included one analysis on the two group means within the higher ranking, another analysis on the two group means within the medium ranking, and the other on the two group means within the lower ranking. Results of the analyses showed that type of instruction influenced the task value of the higher ranking students ($F_{1, 77} = 55.56$, $p = .000$), the medium ranking students ($F_{1, 77} = 117.80$, $p = .000$), as well as the lower ranking students ($F_{1, 77} = 113.02$, $p = .000$). An examination of the interaction plot (see Figure 7) showed that while significant differences existed between the two instruction groups across all three ranking levels, the difference between the lower groups appeared to be relatively greater than the differences of the other two groups.

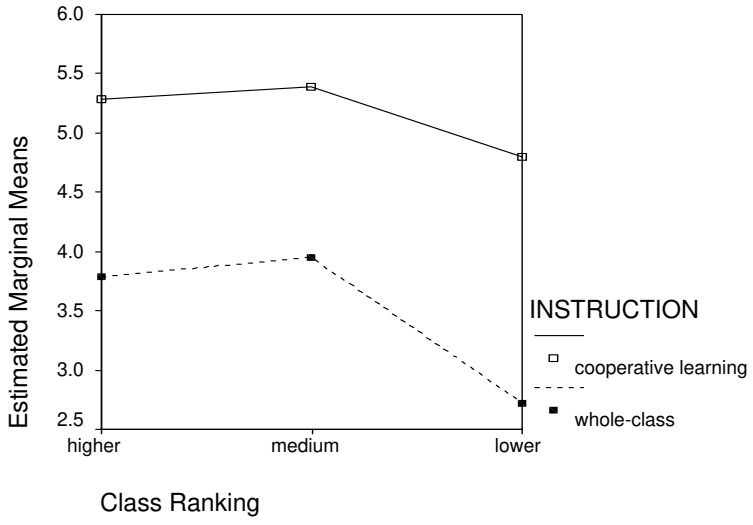

Figure 7. Interaction plot for ranking and instruction on task value.

Exploratory Question B

The second exploratory question was: Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on out-of-class utilization of learning strategies? If so, what is the cause of the interaction? Use of learning strategies was measured in terms of elaboration and peer learning. Again, each scale was analyzed with two, two-way ANCOVAs with instruction (CL vs. WC) and ability level (higher, medium, lower) as the independent variables. The first factorial ANCOVA used the GEPT standard to define ability level; the second used students' relative ranking in class as the operational definition.

Elaboration

Results from the two, two-way ANCOVAs on elaboration (see Table 37) showed no significant interaction effect between GEPT level and instruction ($F_{2, 77} = 1.17$, $p = .315$) while the main effects of GEPT level and instruction were both significant ($F_{2, 77} = 4.13$, $p = .020$ and $F_{1, 77} = 94.03$, $p = .000$ respectively). The results revealed that the effect of instruction on use of elaboration strategies did not depend on which GEPT level was being considered, and vice versa.

On the other hand, significant interaction was found between class ranking and instruction ($F_{2, 77} = 6.54$, $p = .002$). The observed power was excellent at .90, and the effect size was large (eta squared = .15). The statistically significant interaction revealed that the effect of instruction on elaboration depended on which ranking level was being considered.

To examine where the significant ranking by instruction interaction took place, three simple effect analyses were conducted, including one on the two group means

Table 37
Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Elaboration

Source	SS	df	MS	F	p
Ability Defined by GEPT Standard					
Pretest	8.01	1	8.01		
GEPT level	1.95	2	.97	4.13	.020
Instruction	22.16	1	22.16	94.03	.000
GEPT level x Instruction	.55	2	.28	1.17	.315
Error	18.15	77	.24		
Ability Defined by Relative Ranking in Class					
Pretest	5.04	1	5.04		
Ranking	2.51	2	1.26	6.56	.002
Instruction	30.84	1	30.84	161.32	.000
Ranking x Instruction	2.50	2	1.25	6.54	.002
Error	14.72	77	.19		

within the higher ranking, another on the two group means within the medium ranking, and the third on the two group means within the lower ranking.

The simple effect analyses indicated that type of instruction had some bearing on the elaboration strategy use of the higher ranking students ($F_{1, 77} = 57.04$, $p = .000$), medium ranking students ($F_{1, 77} = 39.98$, $p = .000$), and lower ranking students ($F_{1, 77} = 64.99$, $p = .000$). Nevertheless, it is worth noting that in spite of significant differences between the two instruction groups across all three ranking levels, the difference became smaller when it came to the middle-ranking groups (see Figure 8).

Peer Learning

Results from the two, two-way ANCOVAs on peer learning are presented in Table 38. An examination of the results indicated a significant GEPT level by

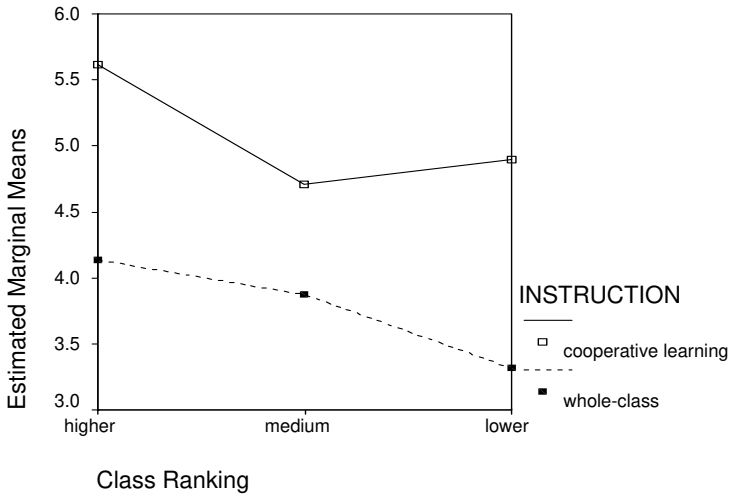

Figure 8. Interaction plot for ranking and instruction on elaboration.

Table 38
Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Peer Learning

Source	SS	df	MS	F	p
Ability Defined by GEPT Standard					
Pretest	1.47	1	1.47		
GEPT level	6.66	2	3.33	15.26	.000
Instruction	35.64	1	35.64	163.46	.000
GEPT level x Instruction	3.58	2	1.79	8.22	.001
Error	16.79	77	.22		
Ability Defined by Relative Ranking in Class					
Pretest	2.36	1	2.36		
Ranking	5.15	2	2.58	11.46	.000
Instruction	38.52	1	38.52	171.36	.000
Ranking x Instruction	2.93	2	1.46	6.51	.002
Error	17.31	77	.23		

instruction interaction effect ($F_{2, 77} = 8.22$, $p = .001$). The observed power was excellent at .95, and the effect size was large (eta squared = .18). Similarly, a significant interaction between class ranking and instruction ($F_{2, 77} = 6.51$, $p = .002$) was found. The observed power was excellent at 1.00, and the effect size was large (eta squared = .15).

To find out where the significant GEPT level by instruction interaction took place, three simple effect analyses were conducted, including one on the two group means within the higher GEPT level, another within the medium GEPT level, and the other within the lower GEPT level. Results of the analyses showed that instruction had an impact on all levels ($F_{1, 77} = 60.65$, $p = .000$, $F_{1, 77} = 70.21$, $p = .000$, and $F_{1, 77} = 55.51$, $p = .000$ for higher, medium, and lower GEPT students, respectively). A scrutiny of the interaction plot (see Figure 9), however, revealed that although significant differences between the two instruction groups were present across all three GEPT levels, the difference at the higher GEPT level was the greatest.

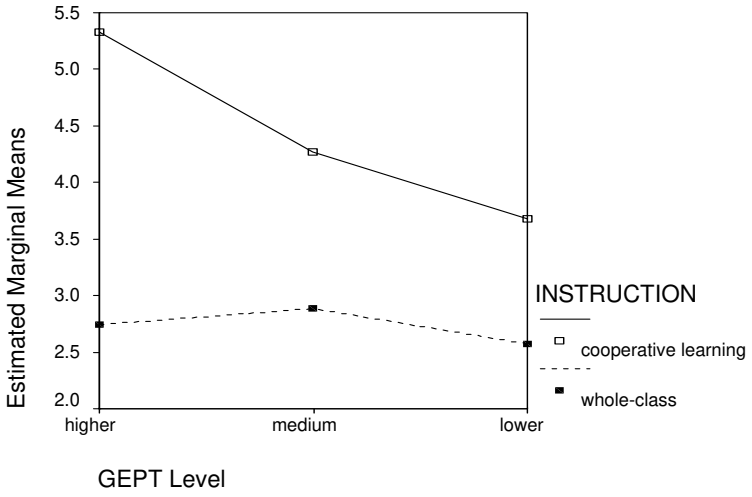

Figure 9. Interaction plot for GEPT level and instruction on peer learning.

To examine where the significant ranking by instruction interaction took place, three simple effect analyses were conducted. Once more, they consisted of an analysis on the two group means within the higher ranking, one on the two group means within the medium ranking, and the other on the two group means within the lower ranking. Results of the analyses indicated that type of instruction influenced all ranking levels of students, including the higher ranking students ($F_{1, 77} = 87.77$, $p = .000$), medium ranking students ($F_{1, 77} = 52.09$, $p = .000$), and lower ranking students ($F_{1, 77} = 41.45$, $p = .000$), on how they used peer learning strategies out of class. A close look at the interaction plot (see Figure 10) showed even though there were significant differences between the two instruction groups across all three ranking levels, the difference was relatively smaller between the two medium ranking groups.

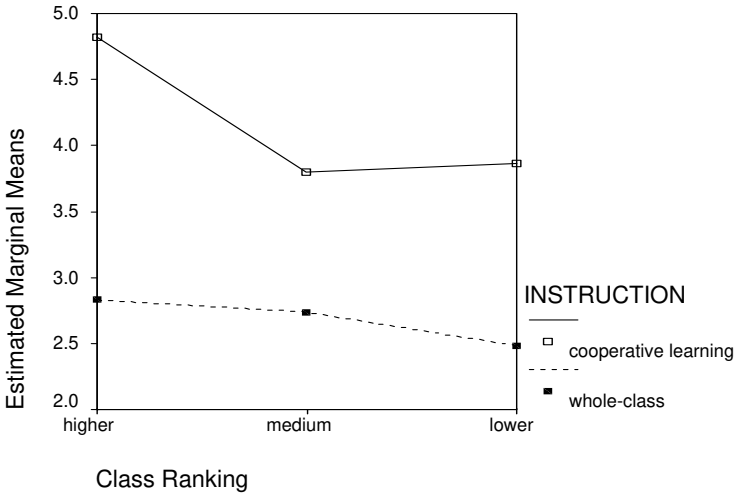

Figure 10. Interaction plot for ranking and instruction on peer learning.

Exploratory Question C

The third exploratory question was: Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on grammar achievement? If so, what is the cause of the interaction? This question was answered with two, two-way ANCOVAs with instruction (CL vs. WC) and ability level (higher, medium, lower) as the independent variables. Once again, the first factorial ANCOVA used the GEPT standard to define ability level; the second used students' relative class ranking as the yardstick for operational definition. Results from the two, two-way ANCOVAs are displayed in Table 39.

Table 39
Factorial ANCOVAs for Interaction Effect of Ability Level and Instruction on Grammar Achievement

Source	SS	df	MS	F	p
Ability Defined by GEPT Standard					
Pretest	4205.44	1	4205.44		
GEPT level	2340.05	2	1170.02	24.22	.000
Instruction	542.17	1	542.17	11.23	.001
GEPT level x Instruction	70.53	2	35.27	.73	.485
Error	3719.13	77	48.3		
Ability Defined by Relative Ranking in Class					
Pretest	2855.15	1	2855.15		
Ranking	1895.46	2	947.73	19.05	.000
Instruction	865.79	1	865.79	17.40	.000
Ranking x Instruction	447.03	2	223.52	4.49	.014
Error	3831.17	77	49.76		

An examination of Table 39 indicated no significant interaction between prior GEPT level and instruction ($F_{2, 77} = .73$, $p = .485$) while the main effects of prior GEPT

level and instruction were both significant ($F_{2, 77} = 24.22$, $p = .000$ and $F_{1, 77} = 11.23$, $p = .001$ respectively). The results showed that the effect of instruction on grammar achievement did not depend on which GEPT level was being considered, and vice versa.

On the other hand, there was a significant interaction effect ($F_{2, 77} = 4.49$, $p = .014$) between ranking and instruction. The observed power was adequate (.75), and the effect size was medium-to-large (eta squared = .10). The statistically significant interaction indicated that the effect of instruction on grammar achievement depended on which prior class ranking level was being taken into account of.

To examine where the significant ranking by instruction interaction occurred, three simple effect analyses were conducted. Like the previous analyses for Exploratory Questions A and B, the simple effect analyses included one analysis on the two group means within the higher ranking, another analysis on the two group means within the medium ranking, and the other on the two group means within the lower ranking. Results of the analyses revealed that, of students who ranked on the top and bottom quartiles, those who received cooperative learning had higher grammar achievement scores than did students who received whole-class instruction ($F_{1, 77} = 14.33$, $p = .000$ and $F_{1, 77} = 5.20$, $p = .025$, respectively). Simple effects at the medium ranking level were not significant, indicating that for students whose relative ranking among classmates were in the middle two quartiles, students in the cooperative group and the whole-class group had similar grammar achievement ($F_{1, 77} = .44$, $p = .510$). The plot in Figure 11 helps visualize the differential effects.

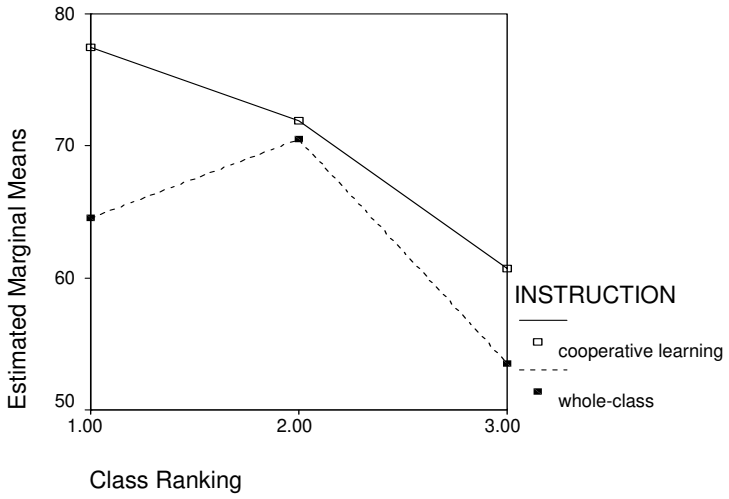

Figure 11. Interaction plot for ranking and instruction on grammar achievement.

Exploratory Question D

The fourth exploratory question was: How does student achievement differ between the cooperative learning group and the whole-class instruction group in terms of different cognitive levels? This question was answered with ANCOVAs, with grammar pretest scores as the covariate, on the grammar posttest scores of various cognitive levels (i.e., the levels of "remember," "understand," "apply," "analyze," "evaluate," and "create"). These levels of cognitive processes are based on Anderson and Krathwohl's (2001) revision of Bloom's taxonomy of educational objectives. (For a detailed discussion of these levels, see the Validity heading in the Grammar Posttest section in Chapter Three.) In the ANCOVA analyses, the proportion of questions answered correctly at each cognitive level was used as the posttest score. The scores are presented in decimals.

In each of the ANCOVAs, all of the assumptions underlying the use of ANCOVA were met. The assumption of normality was not tested as the procedure is robust to its

violation. The assumption of independence of observations was assumed to be true. The covariate (i.e., the pretest) was measured without error. Even though the sample sizes were equal, the homogeneity of variance assumption was tested using Levene's statistic. The observed value was $F_{1, 82} = 3.28$, $p = .074$, indicating equivalency of variances across the groups. The scatterplot and the residual plot were used to test the linearity. The Tolerance and VIF values equaled 1. It was thus concluded that the assumption of linearity was met. In addition, the assumption of homogeneity of regression was tested. The observed value was $F_{14, 50} = 1.37$, $p = .203$ for the "remember" level, $F_{14, 50} = 1.81$, $p = .064$ for the "understand" level, $F_{14, 50} = .31$, $p = .990$ for the "apply" level, $F_{14, 50} = 1.69$, $p = .087$ for the "analyze" level, $F_{14, 50} = 1.44$, $p = .170$ for the "evaluate" level, and $F_{14, 50} = 1.82$, $p = .062$ for the "create" level. The results showed that the lines of the dependent variable that had been regressed on the covariate within each group were parallel, and thus the assumption of homogeneity of regression was satisfied.

An examination of the results indicated no significant difference between the two comparison groups on the lower cognitive levels (i.e., "remember" and "understand" levels) and the medium cognitive levels (i.e., "apply" and "analyze" levels). Among these non-significant results, nevertheless, it is worth noting that while the whole-class group on average scored relatively higher on the "remember" level, the cooperative group was relatively higher on the "understand" level and the two medium cognitive levels (i.e., "apply" and "analyze" levels).

When it came to the higher cognitive levels, the cooperative group's average posttest score at the "evaluate" level (adjusted mean = .58) was significantly higher ($F_{1, 81} = 12.96$, $p = .001$) than that of the whole-class group (adjusted mean = .45). Likewise, the

cooperative group's average posttest score at the "create" level (adjusted mean = .56) were significantly higher ($F_{1, 81} = 11.36$, p = .001) than that of the whole-class group (adjusted mean = .42). Overall there was a pattern showing the higher the cognitive levels went, the more the learners gained from cooperative learning.

The results from the ANCOVAs on the "evaluate" and "create" levels are presented in Table 40. The observed power for the ANCOVA on the "evaluate" level was excellent at .95 and the effect size was large (eta squared = .14). The observed power on the "create" level was excellent at .92 and the effect size was relatively large (eta squared = .12).

Table 40
ANCOVAs on "Evaluate" and "Create" Level Posttest Scores with Grammar Pretest Scores as Covariate

Source	SS	df	MS	F	p
Evaluate					
Pretest	5.20	1	5.20		
Group	.39	1	.39	12.96	.001
Error	2.42	81	.03		
Total	8.00	83			
Create					
Pretest	5.81	1	5.81		
Group	.41	1	.41	11.36	.001
Error	2.94	81	.04		
Total	9.16	83			

Exploratory Question E

The fifth exploratory question was: What are the relationships between prior English level, gender, grammar achievement, task value, self-efficacy, use of elaboration strategies, and use of peer learning strategies? This question was answered with a Pearson's product moment correlation matrix.

Table 41
Correlations among English Level, Gender, Motivation, Strategy Use, and Achievement

	1	2	3	4	5	6	7	8
1. Prior English Level/GEPT	- -							
2. Prior English Level/Rank	.74**	- -						
3. Gender	.06	.12	- -					
4. Grammar achievement	.84**	.88**	.17	- -				
5. Task value	.41**	.62**	.13	.61**	- -			
6. Self-efficacy	.74**	.83**	.12	.87**	.75**	- -		
7. Elaboration	.62**	.69**	.08	.79**	.76**	.87**	- -	
8. Peer learning	.29**	.33**	.10	.42**	.65**	.59**	.76**	- -

Note. $^{**}p < .01$

An examination of the results indicated significant positive relations among all the variables except gender. Significant results were found among prior English level, whether defined by the GEPT standard or by students' relative ranking among peers, grammar achievement, motivation, including task value and self-efficacy, and out-of-class use of learning strategies, including elaboration and peer learning. No relation between gender and other variables was found.

Chapter Summary

The findings of this chapter are summarized in Table 42. When the effect of instruction on motivation was investigated in Research Question One, the results were in favor of the cooperative learning group on each dependent variable. When looking at out-of-class use of learning strategies in Research Question Two, the results again supported the cooperative learning group on both dependent variables. When the effect of instruction on grammar achievement was examined in Research Question Three, cooperative learning once more emerged to be the more effective instructional method.

The results of Research Questions One to Three indicated a consistent pattern in favor of the cooperative learning group. Several exploratory analyses were used to see if this pattern could be more fully explained. While there were two sets of factorial analyses, one using GEPT level and the other using relative class ranking, there was only one interaction effect found for GEPT and instruction. Five interaction effects were found for class ranking and instruction. In all but one case, significant differences were found between the instruction groups across all three levels of English ability. In most cases the difference between the higher ability groups and that between the lower groups were relatively greater than the difference between the middle groups. Additionally, significant differences were found between the instruction groups at the higher (i.e., "create" and "evaluate") cognitive levels but not the middle (i.e., "analyze" and "apply") or lower (i.e., "understand" and "remember") cognitive levels of the grammar achievement test.

In summary, it appears that cooperative learning is more effective than whole-class instruction when considering motivation, use of learning strategies, and grammar achievement. If one looks more carefully at subgroups of different class ranking

Table 42
Summary of Findings

Independent Variables	Dependent Variables	Sig.[a]			
Main Effects					
Instruction (CL vs. WC)	Self-Efficacy	Y(CL)			
	Task Value	Y(CL)			
	Elaboration	Y(CL)			
	Peer Learning	Y(CL)			
	Grammar Achievement	Y(CL)			
	"Create" Level	Y(CL)			
	"Evaluate" Level	Y(CL)			
	"Analyze" Level	N			
	"Apply" Level	N			
	"Understand" Level	N			
	"Remember" Level	N			
Interaction Effects			***Simple Effects*** [b]		
Instruction & Ranking	Self-Efficacy	Y	H-rank CL	vs. H-rank WC	Y(CL)
			M-rank CL	vs. M-rank WC	Y(CL)
			L-rank CL	vs. L-rank WC	Y(CL)
	Task Value	Y	H-rank CL	vs. H-rank WC	Y(CL)
			M-rank CL	vs. M-rank WC	Y(CL)
			L-rank CL	vs. L-rank WC	Y(CL)
	Elaboration	Y	H-rank CL	vs. H-rank WC	Y(CL)
			M-rank CL	vs. M-rank WC	Y(CL)
			L-rank CL	vs. L-rank WC	Y(CL)
	Peer Learning	Y	H-rank CL	vs. H-rank WC	Y(CL)
			M-rank CL	vs. M-rank WC	Y(CL)
			L-rank CL	vs. L-rank WC	Y(CL)
	Grammar Achievement	Y	H-rank CL	vs. H-rank WC	Y(CL)
			M-rank CL	vs. M-rank WC	N
	M-GEPT CL		L-rank CL	vs. L-rank WC	Y(CL)
Instruction & GEPT	Self-Efficacy	N			
	Task Value	N			
	Elaboration	N			
	Peer Learning	Y	H-GEPT CL	vs. H-GEPT WC	Y(CL)
			M-GEPT CL	vs. M-GEPT WC	Y(CL)
			L-GEPT CL	vs. L-GEPT WC	Y(CL)
	Grammar Achievement	N			

Note. Sig. = significance; CL = cooperative learning; WC = whole-class instruction; Y = yes; N = no; H=higher; M = medium; L = lower. [a] In parentheses are instructional types that were found to be significantly more effective. [b] Simple effect analyses were conducted as a result of significant interaction effects listed on the left columns.

levels, cooperative learning facilitates motivational development and strategy utilization of learners across all subgroups, but more so with the higher and lower levels. In addition, cooperative learning facilitates grammar achievement of learners at the higher and lower levels, but not the medium level. As to various cognitive levels that cooperative learning has impact on, it appears that, rather than the lower and medium cognitive levels, cooperative learning facilitates learning at the higher cognitive levels.

The findings from the data analyses in this chapter will be discussed in more depth in the following chapter. In addition, discussion of how the findings contribute to the body of existing literature, as well as their implications for theory, research, and practice are also provided in the final chapter of this dissertation.

CHAPTER FIVE

DISCUSSION

The present study was designed to examine the effects of cooperative learning on students' learning motivation, use of learning strategies beyond class settings, and grammar achievement. Comparisons between cooperative learning and whole-class instruction groups were made with a quasi-experimental research design. The experimental group received the Student Teams Achievement Divisions (STAD) cooperative learning, in which peer tutoring, elaborated explanation, peer modeling, cognitive elaboration, peer practice, and peer assessment and correction were integrated into the instructional strategies as students played important roles in various types of class activities. The control group received the whole-class instruction involving communication-based class activities that were carried out either whole-class (e.g., whole-class discussion) or individually (e.g., journal writing). The duration of the instruction was 12 weeks. Based on the results reported in Chapter Four, cooperative learning was more effective in terms of enhancing learning motivation, use of learning strategies, and student achievement. Discussion of the results will be presented in this chapter according to the major research questions and subsequent exploratory questions. Included within each set of discussion are summary of findings, explanation of findings, and findings in relation to existing literature. The major research questions are:

1. How does motivation differ between the group receiving cooperative learning and the group receiving whole-class instruction?

2. How does utilization of learning strategies beyond class settings differ between the group receiving cooperative learning and the group receiving whole-class instruction?
3. How does grammar achievement differ between the group receiving cooperative learning and the group receiving whole-class instruction?

The exploratory questions are:

A. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on motivation? If so, what is the cause of the interaction?
B. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on out-of-class utilization of learning strategies? If so, what is the cause of the interaction?
C. Is there an interaction effect between instruction (CL vs. WC) and prior English ability level on grammar achievement? If so, what is the cause of the interaction?
D. How does student achievement differ between the cooperative learning group and the whole-class instruction group in terms of different cognitive levels?
E. What are the relationships among prior English level, gender, grammar achievement, task value, self-efficacy, use of elaboration strategies, and out-of-class peer collaboration behaviors?

In addition, implications for theory, research, and practice, delimitations and limitations of the present study, as well as suggestions for future research, are presented in this chapter.

Effects of Cooperative Learning on Motivation

The first research question investigated the effects of cooperative learning on learning motivation. Motivation was measured with the revised MSLQ (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) scales. Results indicated that students who received the cooperative learning treatment displayed higher sense of efficacy in learning and performing and attached higher task value to the coursework than those who received whole-class instruction. More specifically, cooperative learning enhanced learners' performance expectations and their confidence level in mastering learning tasks. Cooperative learning also brought about higher perception on the course material and task in terms of its interest, significance, and usefulness. The findings support the hypothesis that the students in the cooperative learning group would show higher motivation than those in the whole-class group. The findings also converge with those of previous research on the benefits of cooperative learning to second language students in the affective domain (e.g., Ghaith, 2002, 2003b; Ghaith & Bouzeineddine, 2003; Liang, 2002).

The higher level of self-efficacy displayed by the cooperative learning learners can be explained in light of the expectancy theories, including Weiner's (2000) attribution theory, Bandura's (1993) self-efficacy theory, and Covington's (1992) self-worth theory. Learning motivation, according to Weiner (2000), is subject to learners' attributions of past performance, either success or failure. Learners who attribute their past performance to stable, constant, and hence uncontrollable factors, e.g., inherent ability, tend to give up more easily on a task and develop less motivation for learning than those who attribute their performance to unstable, temporary, and hence

controllable factors, e.g., level of effort. The structure of the STAD cooperative learning method employed in the present study created a situation in which learners were evaluated based upon the level of personal improvement. This feature, which is also called "equal opportunities for success," allowed the cooperative learners to perceive success as something attainable by effort rather than something that could fall beyond reach due to inherent ability or keen competition. Lower achievers might find this feature of STAD motivating as they were given chances to succeed on their own terms instead of having to be constantly compared with higher achievers. Meanwhile, higher achievers might also perceive more in control of their learning because, rather than competing intensively against other higher achievers, their objective was to excel themselves.

The advantage of the self-improvement, equal-opportunity feature in STAD is also supported by Bandura's (1993) self-efficacy theory, which asserts that learners are more apt to assess their ability by their personal improvement if they perceive competence as acquired skills. The self-improvement feature in the experimental group geared away from the traditional ranking system, focused on personal development, and helped learners at different performance levels to identify competence as acquired. If the experimental students stumbled upon difficult tasks, this feature could allow them to examine the processes such as effort exerted and strategies used and to keep a task-diagnostic focus and concentrate on how to perform successfully instead of maintaining a self-diagnostic focus and falling as an easy victim to stress. In case of disappointing performances, it would be easier for them to recover their sense of self-efficacy because failure mostly meant inadequate endeavor or insufficient knowledge and capacities that were attainable. Perceiving themselves more in control of their own

learning by perceiving success as the outcome of hard working and effective strategy use, they could thus become more motivated and ready to face challenges.

Based on Covington's (1992) self-worth theory, in order to maintain a sense of self-worth and self-control, learners with low self-esteem tend to shy away from working hard so that they can attribute failure to the level of effort put forth (i.e., a controllable factor according to the attribution theory mentioned above). In this regard, STAD created a condition in which the experimental students at various performance levels need not worry about competing with others; they only needed to exert effort so that they could be better than how they had been. In other words, success became more within reach. When success became feasible by way of effort, they did not have to shy away from working hard to save their sense of self-worth because now the level of effort and even the possibility of success were both controlled in their own hands.

In addition to higher motivation in terms of self-efficacy, the cooperative learning group in the present study also demonstrated higher task value than the whole-class group. As discussed in the review of literature in Chapter Two, Eccles and her colleagues (Eccles, 1987; Eccles & Wigfield, 1995, 2002; Eccles (Parsons) et al., 1983) have identified four kinds of task values: attainment value, intrinsic value, utility value, and cost. The task value scale employed in the present study basically measured the first three types of task values. Overall the cooperative learners perceived a stronger relationship between course tasks and their current or future goals, attached higher personal importance on performing well upon course learning tasks, and experienced more enjoyment and pleasure when carrying out course-related tasks.

Higher task value among the cooperative learners can be discussed in light of certain cooperative learning elements in relation to the model of triadic reciprocality (Bandura, 1986) and goal setting theory (Locke & Latham, 1990). First, peer modeling was a recurring event in the experimental group as a result of the heterogeneous grouping and the positive interdependence features of cooperative learning (see Chapter Two for a detailed description). Based on the model of triadic reciprocality, modeling and subsequent social persuasion could shape, lead, and transform the cooperative learners' thoughts and feelings, enhance stimulation, and arouse their emotions. According to the goal setting theory, the cooperative learners could obtain higher personal goal setting and higher goal commitment by observing a higher-performing role model.

Secondly, when cooperative learning is structured and implemented properly, encouragement and feedback among peers occur because of the individual accountability and the positive interdependence features of cooperative learning. During class observations that aimed to ensure treatment fidelity, the researcher of the present study observed verbal encouragement exchanged among cooperative learners. She also observed various types of performance feedback, including verbal assessment on academic performance and verbal correction on specific tasks among peers, calculation of individual improvement points (see Table 25 in Chapter Three), and written peer evaluation on social and cooperative performance. Anchored on the goal setting theory (Locke & Latham, 1990), the encouragement and performance feedback taking place in the cooperative group could have facilitated higher goal setting and thus higher task value.

With reference to past studies on how cooperative learning affects Taiwanese EFL students' learning motivation (Chen, 1998; Chu, 1996; Liang, 2002; Lo, 1998; Wang, 2001; Wei, 1996; see Chapter Two for a detailed review), the present study contributes to the existing literature in several ways. First of all, compared to the studies that used college students and questionnaires (Chu, 1996; Lo, 1998; Wang, 2001; Wei, 1996), the present study extends the findings by specifically looking into task value and self-efficacy, and by employing a different cooperative learning method, i.e., STAD. The methods involved in the other studies included Jigsaw, Group Investigation, and Learning Together. Secondly, in comparison with some of the above-mentioned studies that also used a quasi-experimental research design, the present study extends the findings by enriching the whole-class instructional program with communicative activities to ensure that the control group would also receive quality teaching, by extending the length of experiment to enhance validity, and by separating the roles of the researcher and the classroom instructor to increase objectivity. Thirdly, the present study utilized a measurement tool with strong validity and reliability to measure student motivation. Finally, in comparison with the studies with a similar research design and the same cooperative learning method (Chen, 1999; Liang, 2002), the present study extends the findings by focusing on a different population (i.e., college versus junior high and high school EFL students) and by employing a different measurement tool to understand student motivation.

Effects of Cooperative Learning on Utilization of Learning Strategies

The second research question investigated the effects of cooperative learning on utilization of learning strategies beyond class settings. Two revised scales from the

MSLQ (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) were employed to understand how the treatment impacted the students' out-of-class learning strategy usage, specifically elaboration and peer collaboration strategies. It was hypothesized the cooperative learners would demonstrate more frequent use of the learning strategies than those in the whole-class group. Results from Chapter Four converge with the hypothesis. Cooperative learning has enhanced the experimental learners' use of learning strategies beyond the formal class setting. When preparing for the course, the experimental learners turned out to use more elaboration skills, including paraphrasing, summarizing, and synthesizing, and more out-of-class peer collaboration than the control learners.

To discuss possible reasons that attributed to the cooperative learners' more frequent use of elaboration strategies outside the class context when they were preparing for the course, one needs to first have a good grasp of what took place inside the experimental class via the lenses of the cognitive elaboration theories (Dansereau, 1988; Walling, 1987) and the cooperative learning model (Slavin, 1995). Based on the group-goal structure of cooperative learning (see Figure 2 in Chapter Two), the success of group work in the experimental class depended immensely on peer discussion and tutoring, that, of course, involved a great deal of presentations of thoughts as well as explanations on the subject matters. The learners thus obtained opportunities to articulate and elaborate their preliminary, immature thoughts. Ideas could be cultivated from vague to concrete and from premature to refined during the explanation and elaboration processes. In the meantime, the articulating and elaborating processes could result in active processing of information, cognitive restructuring, and reprocessing of thought (Dansereau, 1988), which, in turn, aided skill development on paraphrasing, summarizing,

and synthesizing. It appears that the cooperative learners' elaboration skills could have been practiced and sharpened through peer practice, peer explanation, and peer modeling. It was likely that the frequent use of elaboration strategies within cooperative groups during class time allowed the learners to transfer the elaboration skills to individual settings beyond the classroom context. The above inference as to how cooperative learning could facilitate elaboration skills for self-study is summarized in Figure 12.

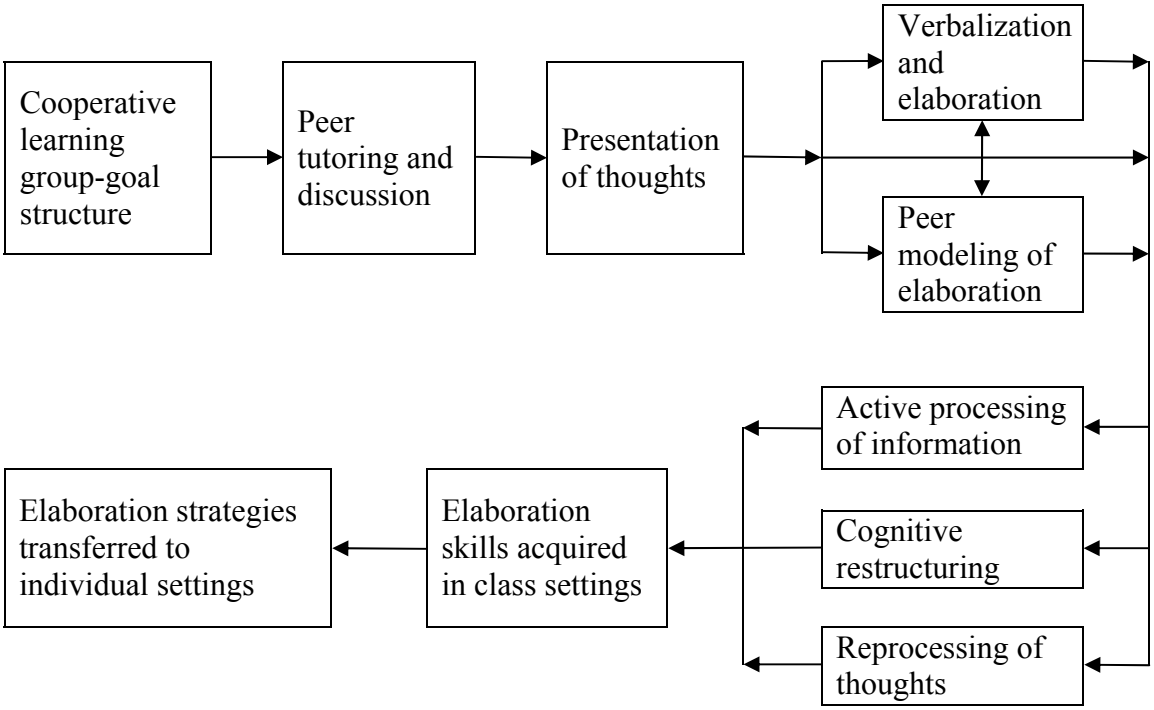

Figure 12. Possible process of how cooperative learning facilitates elaboration skills for self-study.

In addition to more frequent use of elaboration strategies, the cooperative learners have also utilized more peer collaboration strategies beyond the class setting than the

whole-class learners. One possible reason was that the study habits of the cooperative learners were molded by their class activities. The structure of the cooperative learning instruction made peer collaboration a routine. As stated by Bernard (1926) and Robinson (1995), human beings are habit forming animals. After working within groups for an extensive period of time over the semester, it was possible that the learners formed the habit of studying together with peers. It therefore appears reasonable to attribute the cooperative learners' more frequent use of out-of-class peer learning strategies to habit.

Another possible explanation for the cooperative learners' enhanced out-of-class collaboration goes deeper beneath habit forming to the learners' conscious decision making. As indicated in Chapter Three, positive interdependence and individual accountability were not just slogans in the cooperative class. There were actually systematic ways to hold the learners accountable for their teammates' academic development. Each person's success depended upon his or her teammates' academic growth. Once the learners realized that none of them could succeed without the other teammates' achievement, they would want to work together—either to help others or seek others' help, or both. If the experimental learners did not assist each other to accomplish academic enhancement, they would not be able to accomplish their group goals. It was very likely that this "sink or swim together" (Johnson, Johnson, & Smith, 1991, p. 16) feature had made the cooperative learners extend their use of peer collaboration beyond classroom contexts.

In comparison with other studies on the effects of cooperative learning in the EFL contexts, the current study adds to the existing literature by extending the findings to students' strategy use beyond the classroom. Strategy use in out-of-class self-study

settings has not been much researched. To researchers the findings of the present study on use of learning strategies contribute to a new dimension on the effects of cooperative learning. To practitioners the findings will help develop curricula to better fit their specific teaching aims through understanding how cooperative learning changes and enhances students' out-of-class study behaviors.

Effects of Cooperative Learning on Grammar Achievement

The third research question investigated the effects of cooperative learning on grammar achievement. The findings in Chapter Four indicated that the cooperative learners on average demonstrated higher grammar achievement than the whole-class learners. The findings support the hypothesis stated in Chapter Three that the cooperative learning students would display higher grammar achievement than the control group students. The findings are also congruent with the past research findings on the positive impact of cooperative learning on academic achievement in a wide range of subject areas (see Johnson & Johnson, 1989 and Slavin, 1995), as well as with those in the ESL and EFL fields (e.g., Calderon, Hertz-Lazarowitz, & Slavin, 1998; Liang, 2002; Ghaith, 2003a, 2004).

Liang (2002), in a discussion on how cooperative learning could have positively affected the language competence of her experimental learners, emphasized three factors: the increase in student talk for academic and social purposes, the incentive structure of positive interdependence, and the supportive and communicative learning environment. Survey studies conducted by Ghaith (2002, 2003b) indicated that the cooperative learners felt more academic and personal support from their peers and teachers, more class cohesion and fairness of grading, and less school alienation. Another survey study

(Ghaith, 2001) suggested clear cooperative structure and lucid guidance as the possible reasons for positive cooperative learning effects. The findings of the present study suggest that these factors may also have played a role in enhancing the experimental learners' grammar achievement. Other possible reasons can be explained in light of the following motivational theories, social cognitive theories, and cognitive elaboration theories.

First, according to Locke and Latham's (1990) goal setting theory, human behaviors, which, of course, include learning behaviors, are regulated by goals. Factors influencing the level of learners' goal setting and goal commitment include whether group goals, encouragement, and feedback exist in the learning situation. Having group goals on top of personal goals leads to stronger goal commitment to the personal goals than having merely personal goals; giving encouragement and performance feedback enhances level of goal setting. All these three factors—group goals, encouragement, and performance feedback—were essential parts of the experimental curriculum. Like all other cooperative learning methods (refer to Figure 2 in Chapter Two), the starting point of the experimental program was having group goals, which would motivate learners to offer both academic and psychological support, including encouragement and feedback, to each other. In the cooperative learning group, feedback was rendered not only through peer assessment and correction but also through routine calculation of individual improvement points (IIP) and group average improvement points (GAIP). The design of the calculation worksheets (refer to Tables 25 and 26 in Chapter Three) has made it easier for the cooperative learners to monitor both their individual and group progresses.

Secondly, based on Weiner's (2000) attribution theory, Bandura's (1993) self-efficacy theory, and Covington's (1992) self-worth theory, it was possible that the structure of the STAD method, especially its equal-opportunity feature, enabled the cooperative learners to perceive competence as acquired, which in turn made them believe in the worth of themselves, effort making, and constant self-improvement.

The third possible reason for higher grammar achievement of the cooperative learning group can be attributed to enhanced motivation, which has been explored at length earlier in this chapter when discussing the results of Research Question One regarding effects of cooperative learning on learning motivation. Hence details on that are not going to be repeated. The point to be made here is that the enhancement in perception and belief caused by the implementation of cooperative learning very likely would positively transform the cooperative learners' learning behaviors and consequently produced better academic outcome.

The fourth possible explanation for higher grammar achievement of the cooperative learning group can be ascribed to triadic reciprocality (Bandura, 1986) among environmental factors, personal and cognitive factors, and learning behaviors (refer to Figure 3 in Chapter Two). Through verbal modeling of thought process and social persuasion, learners' thoughts can be shaped, directed, and modified; stimulation can be enhanced; and learning can be facilitated. Bandura reminded that the effect of triadic reciprocality takes time, but since the experiment of the present study was implemented over a course of 12 weeks, there had likely been sufficient time to allow peer modeling to begin exerting impact on the cognitive skill development (i.e., grammar achievement) of the subjects.

The fifth possible reason can be viewed in light of Vygotsky's (1978) zone of proximal development (ZPD) premise and Piaget's (1964) social transmission theory. Based on these social cognitive theories, new information becomes accessible to learners only when they already have a structure which allows them to assimilate the information. When the learners have such a structure in a learning situation, they are actually in what Vygotsky called the zone of proximal development. A good way to increase learners' ZPD activities is to let them work in peer groups, in which learners can help one another move to the next level of development. It seems that peer explanation in the cooperative learning group of the present study allowed the learners to work closely within one another's ZPD and to receive elucidations that were presented to them in a simpler and more comprehensible way than if they had been presented by one of a very different intellectual age, such as the instructor in the control group.

The sixth possible explanation for cooperative learning's positive effect on grammar achievement can be discussed in light of the equilibration theory (Piaget, 1932, 1950, 1964). Piaget (1964) argued that all advancement comprises "momentary conflicts and incompatibilities which must be overcome to reach a higher level of equilibrium" (p. 19). To be more specific, learners first use their existing schemes to make sense of the world; when they notice that the current schemes can no longer encapsulate the reality of the external world, they either amend the current schemes or construct new ones to restore balance. The equilibration process facilitates learning and development (1950).

While equilibration process can also take place with learners working individually (Piaget, 1932), such as those in the control group, it is reasonable to assume that, through peer explanation, peer tutoring, and group negotiation, the experimental learners had

more opportunities for equilibration to occur than the control learners who worked alone. Even when whole-class discussion took place in the control group, the effect of equilibration was probably discounted due to the nature of discussion and the learners' anxiety level. To start with, only a limited number of students at a time were able to be engaged in a whole-class intellectual conversation. While a few highly motivated students might go through the equilibration process as audiences, most others would easily allow themselves to fall into passive listeners and thus received very little equilibration effect. What's more, Taiwanese students were prone to high anxiety when speaking in front of the whole class (Babcock, 1993; Liao & Chang, 2003; Liao & Hsueh, 2005). The higher level of anxiety the learners experienced, the less likely equilibration would have an effect. Quite the opposite, small group activities in the cooperative group appear to have given students a more private and psychologically safer setting to express themselves and negotiate their thoughts. Because group sizes were small, instead of easily shirking participation, learners had to participate in academic dialogues and thus had opportunities to go through the equilibration process as described by Piaget (1950, 1964).

Finally, in light of cognitive elaboration theories (Dansereau, 1988; Snowman & Biehler, 2005; Van Boxtel, 2000, cited in Veenman, Denessen, Van Den Akker, & Van Der Rijt, 2005; Walling, 1987), opportunities to construct explanations and elaborations within groups could represent a significant arbitrator of positive cooperative learning effects on grammar achievement. In a traditional lecture-type course where the instructor hand-feeds information, students may be granted less time and fewer chances to develop their ideas from preliminary to sophisticated. Certain thoughts may come to the learners'

minds during lectures; some of them may be full of potential and worth exploring. Yet due to the nature of lecturing, they may come and go in a flash and fade away before they can sprout and mature. Not enough time is granted for reflection and digestion of new information. Conversely, through interacting and negotiating with their teammates, the cooperative learners in the present study were able to elaborate their thinking and actively process information. Compared to the control learners, they had more time to reflect upon the subject matter, raise their awareness, structure and restructure knowledge, differentiate information received, fine-tune their thoughts, and expand their knowledge base (Dansereau, 1988; Van Boxtel, 2000, cited in Veenman, Denessen, Van Den Akker, & Van Der Rijt, 2005). As a result of these cognitive activities, they were able to learn more effectively, retain information longer, and thus perform better on the grammar achievement test (Dansereau, 1988; Snowman & Biehler, 2005).

With reference to previous studies on how cooperative learning affects Taiwanese EFL learners' academic performance in English (Chang, 1995; Chen, 1999; Liang, 2002; Wang, 2001), the present study contributes to the existing literature in numerous aspects. First, it improves on previous research by having a quasi-experimental pretest-posttest comparison group research design (versus a one-sample repeated-measures design or a quasi-experimental design without a pretest). This research design not only avoids carryover effects and progressive error but also controls potential pretreatment differences between the comparative groups. Secondly, care was taken to ensure valid and reliable instrumentation for both the pretest and the posttest. Thirdly, the present study extends the findings by having an enriched control program. Instead of comparing cooperative learning to traditional lecture courses, which have been found to be

uninspiring by many EFL learners in Taiwan (Lai, 2001; Su, 2003), context-rich instruction and communicative activities were added to the whole-class control curriculum. Fourthly, this study improves on past investigations by employing ANCOVA (versus multiple t tests or ANOVA) to reduce Type I error and the estimate of random or error variance in the dependent measure. Finally, the study expands the horizon of the related literature by reporting practical significance in addition to statistical significance. With this extra information, practitioners and school administrators will be better informed as to how practical and beneficial it is to implement cooperative learning in their classrooms or institutions.

Interaction Effects of Instruction and Prior English Level

The findings of the three major research questions consistently showed cooperative learning as more favorable than whole-class instruction. As a result of the consistent pattern, several exploratory questions were posed to see if the pattern could be more fully investigated. Exploratory Questions A to C examined whether there was an interaction effect between type of instruction and prior English level on motivation, strategy use, and grammar achievement. A very interesting finding on these exploratory questions is that, while the effects of instruction in most cases did not depend on which GEPT level was being considered, there were differential effects as to which class ranking was being taken into account of. This implies that the effects of cooperative learning depended on students' ranking in relation to peers rather than on objective performing levels identified by measures such as standardized tests. It is also reasonable to assume that the effects depended on how students perceived their relative ranking and subsequent role-taking as a result of the perception.

Despite the fact that cooperative learning enhanced motivation and strategy use for students across all class ranking levels, the effects were relatively greater for the higher and lower ranking students. In the case of grammar achievement, the higher and lower ranking students benefited from cooperative learning, but the middle ranking cooperative learners only performed at the same level as the middle ranking whole-class learners.

Among numerous reasons that could contribute to the differential effects in favor of cooperative learning, the higher ranking students might have particularly profited from their natural explainer roles. As for lower ranking students, after being under-achievers for probably most of their school life, the "equal opportunity" "self-improvement" structure of the cooperative learning method could have helped them perceive learning in a different light. Specifically, they might come to see competence as acquired skills and that they too had power to make improvement and reach success.

The fact that the middle ranking learners received no effect on grammar achievement when compared to the whole-class learners could be discussed in light of past research findings regarding medium ability students' performance in cooperative learning groups. In the past literature, medium ability learners were found to perform better in homogeneous small groups instead of heterogeneous small groups (Abrami, Chambers, Lou, Poulsen, & Spence, 1999; Lou, Abrami, & Spence, 2000). Webb and Palincsar (1996) also reported that when working with high ability peers in four-tiered ability groups, medium ability learners missed many opportunities to construct explanations. Based on these research findings, it appears that the four-tiered ability group structure in the present study could have caused the middle ranking learners to play

more of listener roles than explainer roles within their small groups. Seeing the presence of more-capable peers in their teams, they might have shrunk from responsibilities and allowed themselves to fall into the passive roles of listeners, similar to what they might have normally done in whole-class lecture settings.

Nevertheless, while the middle ranking learners in the cooperative learning group did not outperform the middle ranking learners in the whole-class instruction group in English grammar achievement, they were able to perform at the same level. In addition, they demonstrated higher motivation and more frequent utilization of learning strategies. If we also take into account of the higher and lower ranking cooperative learners, who not only displayed higher grammar achievement but also higher motivation and better utilization of learning strategies in comparison to their counterparts in the whole-class instruction group, on the whole cooperative learning appears to be a more promising pedagogy.

Effects of Cooperative Learning on Performance at Different Cognitive Levels

Exploratory Question D was an extension of Research Question Three. It investigated the effects of cooperative learning on English grammar achievement at different cognitive levels. The findings in Chapter Four indicated that while the cooperative learners performed at a similar level as the whole-class learners in terms of lower (i.e., "understand" and "remember") and middle (i.e., "analyze" and "apply") cognitive question items, cooperative learning enhanced the learners' performance when question items of higher cognitive (i.e., "create" and "evaluate") levels were considered. There was a pattern where the higher the cognitive level became, the more effective

cooperative learning was than whole-class instruction on EFL learners' English grammar achievement.

The present study makes valuable contributions to existing literature as it appears to be the first study systematically analyzing how cooperative learning impacts EFL learners at six different cognitive levels. The findings also have some bearing on how practitioners can make curriculum decisions. If the objective of a curriculum is for students to retrieve pertinent information, to understand instructional messages, or to use the acquired knowledge to a new situation, both cooperative learning and whole-class instruction can be considered. However, if the objective is to facilitate higher-level learning, including for students to check and critique information based on imposed criteria, as well as to create by putting elements together to form a coherent or functional whole, cooperative learning is more favorable.

Relations among Grammar Achievement, Motivation, Strategy Use, and More

Exploratory Question E examined the relationship among grammar achievement, motivation (including self-efficacy and task value), learning strategy use (including elaboration and peer learning), prior English level (including GEPT level and relative class ranking), and gender. The findings in Chapter Four showed that all the above factors were closely related to each other with the exception of gender. Therefore, in order to maximize the effects of instruction, it would certainly help to consider all the closely-knitted factors when designing English language programs for students in Taiwan.

The finding that elaboration strategy utilization positively related to academic achievement converges with the results of past investigations, including those conducted

by Veenman, Denessen, Van Den Akker, and Van Der Rijt (2005), Webb and Farivar (1999), and Webb and Palincsar (1996). Additionally, in their recent study Veenman and his colleagues (2005) suggested a need for further research to determine the relations between different affective-motivational variables and help-seeking or help-giving behaviors. The present study contributes to the existing literature by presenting the findings of positive correlations among self-efficacy, task value, and use of peer collaboration.

Implications of the Present Study

The findings of the present study have improved our understanding of the effects of cooperative learning in a Taiwanese EFL context. Implications of the findings can be discussed in terms of theory, research, and practice.

Theoretical Implications

With respect to theory, Slavin's (1995) cooperative learning model (see Figure 2 in Chapter Two) on which the present study was based is the prevailing theoretical model in the research field of cooperative learning. The model postulates a chain reaction of group goals enhancing motivation, which facilitates cognitive advancement in a social context, which in turn promotes higher achievement. The results of the study are consistent with the model. Cooperative learners in the experimental group showed higher motivation, more learning strategy utilization, and enhanced English grammar achievement. It appears that cooperative learning facilitates learning not only because it stimulates learners with mutual objectives but also because it further places learners in a social framework, which provides an arena for cognitive growth through elaborated

explanations, peer tutoring, peer modeling, cognitive elaboration, peer practice, peer assessment, and peer correction.

Slavin's (1995) model of cooperative learning is sound, yet if related motivational and social cognitive theories can be incorporated into the model, it would make the model even more robust and further advance the understanding of cooperative learning. Motivational theories that can be integrated into the model include attribution theory (Weiner, 2000), self-efficacy theory (Bandura, 1993), self-worth theory (Covington, 1992), and goal setting theory (Locke & Latham, 1990). Social cognitive theories that can be incorporated include zone of proximal development (Vygotsky, 1978), social transmission theory (Piaget, 1964), equilibration theory (Piaget, 1932, 1950, 1964), and model of triadic reciprocality (Bandura, 1986).

Methodological Implications

With respect to research methodology, this study conducted a number of exploratory analyses trying to explain more fully the results of the major research questions. If the investigation had been concluded with the major research questions, our understanding of cooperative learning from the study would have been limited to the differential instructional effects between cooperative learning and whole-class instruction. Conversely, with further exploration through interaction effects and simple effect analyses on instruction and students' relative class ranking, we come to understand that the effects of cooperative learning actually depend on the level of class ranking being considered. In other words, by taking into consideration student characteristics and by utilizing multiple sets of data analysis procedures, we are able to go beyond the differential effects of cooperative learning and get a clear picture of how cooperative

learning and students' relative ranking interplay. In addition, by examining different cognitive levels within the English grammar achievement test and by conducting additional analyses beyond the major research questions, we are able to grasp which cognitive levels cooperative learning affects. The lesson to be learned from these findings is that, in order to understand the essence of a phenomenon, a researcher needs to take a close look at his or her subjects, measurement tools, and data from multiple angles, and then he or she needs to be willing to experiment with diverse analysis procedures to see if the analyses add any new dimension to the phenomenon.

Pedagogical Implications

Based on the findings of the present study, it appears that, in order to enhance learning motivation, instruction needs to be tailored to help learners perceive competence as acquired skills and to enhance their sense of control over learning tasks. In order for EFL learners to perceive competence as attainable through efforts and to make them believe their power in making a difference, allowing them to make improvement against their own past performance rather than against their classmates seems a reasonable solution. In cooperative learning this approach is known as "equal opportunities for success." In order to take advantage of this feature, EFL practitioners could consider the effects of a number of cooperative learning methods, such as Student Teams Achievement Divisions (STAD), Teams-Games-Tournament (TGT), Team-Assisted Instruction (TAI), Cooperative Integrated Reading and Composition (CIRC), and Jigsaw II. (See the Definition of Terms section in Chapter One for brief descriptions of these cooperative learning methods).

In an earlier discussion on possible reasons that could enhance cooperative learners' peer collaboration strategy use beyond the class context, the researcher of the present study postulated that the increase in strategy use could be ascribed to students' conscious decision making, that students understood they must work collectively toward the group goals so that they could succeed as a team. This postulation, nevertheless, was based on the premise that positive interdependence and individual accountability had already existed. Practitioners need to be aware that cooperative learning will not be effective without these two elements. In point of fact, cooperative learning will no longer be true cooperative learning without these elements.

If practitioners do choose cooperative learning as the instructional approach, measures need to be taken to ensure positive interdependence and individual accountability. To ensure positive interdependence, having group goals is an indispensable factor. Other possible measures include resource interdependence and role interdependence. To ensure individual accountability, recommended methods include small group size, peer tutoring, and individual tests, all of which are essential parts of STAD.

In addition, during the class observations that were intended to ensure treatment fidelity of the present study, the researcher noticed the instructor adopting two ways to ensure individual accountability. First, she observed group process and noted the frequency of each student's participation. Second, she randomly called on a student to orally present the group's work in front of the whole class. What the instructor did in class was also recommended by Johnson, Johnson, and Smith (1991), who further suggested assigning a checker in each group to periodically check the team members'

comprehension by asking them to give an explanation of what had been learned. Based on the class observation and the findings of the study, the above methods can provide practitioners systematic ways to ensure individual accountability and positive interdependence. Additional methods can be found in Webb and Farivar (1994), Webb, Farivar, and Mastergeorge (2002), and Webb and Mastergeorge (2003).

The findings of the present study show that, while cooperative learning enhances learning motivation and out-of-class learning strategy use for students across all class ranking levels, it promotes higher grammar achievement for students of the top and bottom quartiles but not for those of the middle two quartiles. For the middle achievers, it does not seem to make a difference in terms of grammar achievement whether they receive cooperative learning or whole-class instruction. If a practitioner's sole objective is to enhance the middle achievers' grammar performance, cooperative learning might be costly as it requires much preparation beforehand and close monitoring during the group process. However, if a practitioner also considers enhancing learning motivation and strategy use of the middle achievers, or if the cognitive or motivational development of the higher and lower achievers is also part of the consideration, cooperative learning is preferred.

Finally, the present study yields some pedagogical implication for practitioners who choose traditional lecture-style instruction as their teaching method. Findings from the present study suggest that one possible reason for the cooperative learners' better academic performance can be ascribed to zone of proximal development activities (or, in an SLA term, "$i + 1$" activities) within peer groups. Therefore, even if lecturing is chosen over cooperative learning based on cost effectiveness or any other reason, the instructors

should still try to teach within students' zone of proximal development. To be exact, the instructors must recognize there is likely an intellectual gap, which could be relatively large, between them and their students on a specific subject matter. The instructors should try to understand where the students' current level is and teach from there. They need to be observant, understanding, and flexible.

Delimitations and Limitations

Several delimitations and limitations of the study have been identified. According to Creswell (1994, 2002), delimitations address how the researcher has purposefully chosen to narrow the extent of the study; limitations identify problems that are beyond the control of the study.

Delimitations

The most salient delimitation of this study is the independent variable. Among various cooperative learning methods, the researcher has chosen to use STAD for the operational definition of cooperative learning. The next delimitation is that the generalization of the results will be delimited to private colleges and universities in Taiwan that have similar entrance scores. Furthermore, because the experiment has been conducted during the regular semester for the duration of 12 weeks, the researcher has chosen to delimit the study to the generalization of results to regular semesters only, not during other situations such as a mini-session or an intensive English language program (IELP). In addition, in the present study the type of measurement for student grammar achievement was a paper-and-pencil test. Grammar achievement, therefore, has been delimited to be performance in a written context instead of an oral communication context.

Limitations

In interpreting the findings of this study, some possible limitations should be recognized. First, the data were collected on only one occasion at pretest and one occasion at posttest. This possibly limits the availability of data that could shed light on the differential effects of cooperative learning over an extended period of time.

Secondly, motivation and use of learning strategies were measured with a self-report questionnaire. It is difficult to discern if the measures accurately represent the subjects' true attitudes and actual use of strategies. Nevertheless, considering the significance of attitudes and beliefs in behavioral science (Ajzen, 1989; Fishbein & Ajzen, 1975), the findings of the present study still present important information on understanding the effects of cooperative learning.

Thirdly, in the present study motivation and strategy use were each measured with two scales. There are other dimensions of motivation and strategy use that future research could explore. This will be discussed in more detail in the Suggestions for Future Research section.

Fourthly, in terms of task value, four types of task values have been identified by Eccles and her colleagues (Eccles, 1987; Eccles & Wigfield, 1995, 2002; Eccles (Parsons) et al., 1983), namely intrinsic value, attainment value, utility value, and cost. The task value scale employed in the present study basically measured the first three types but not the fourth. Thus, future research should involve attempts to develop scales that would take the fourth type of task value into consideration.

Suggestions for Future Research

The following suggestions for future research have emerged as a result of the present study. Regarding independent variables, the current study has chosen to investigate the differential effects of cooperative learning and whole-class instruction. It is worth including the combination of these two types of instruction as an additional independent variable in a future study.

With respect to dependent variables, there are many new dimensions of the current ones that warrant further exploring. For example, motivation can be investigated from a cost-of-task dimension. Cost of task was identified by Eccles and her colleagues (Eccles, 1987; Eccles & Wigfield, 1995, 2002; Eccles (Parsons) et al., 1983) as one type of task values, but it was not in the scope of the present study.

Motivation can also be investigated from an extrinsic-versus-intrinsic motivation dimension. Whereas some critics of cooperative learning (e.g., Kohn, 1991a, 1991b) contend that extrinsic motivation prompted by cooperative learning can deteriorate intrinsic motivation, advocates of cooperative learning and the self-determination theory (e.g., Deci & Ryan, 1985; Deci, Vallerand, Pelletier, & Ryan, 1991; Slavin, 1991a; Swezey, Meltzer, & Salas, 1994; Vallerand, 1997) believe otherwise. They have argued that extrinsic motivation can facilitate intrinsic motivation. To be more specific, they have maintained that a learner's more controlled and extrinsic motivation can be internalized in a social context and become more self-determined and intrinsic motivation. In the present study, the task value scale has been utilized to measure the degree to which the learners perceive the course material and task in terms of interest, significance, and usefulness. Although the items on the scale have touched on various levels of extrinsic

and intrinsic motivation on the internalization continuum proposed by Deci and Ryan (see Figure 1 in Chapter Two for a graphic representation of the continuum), the measurement has given a composite score instead. It is recommended further studies be conducted employing scales that can independently measure different levels of extrinsic and intrinsic motivation so as to examine the relations between different types of motivation and cooperative learning. The researcher of the current study strongly suggests the measure be administered at various points during the studies to grasp a better understanding of the motivation internalization process if there is any.

Other dimensions of motivation that could be further investigated include but not limit to goal orientation and control of learning beliefs. These aspects can be measured with the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991, 1993), although some revision might be required to obtain more robust reliability.

In addition to motivation, use of learning strategies can also be explored more fully in future studies. In a recent study on effects of cooperative learning in a mathematics course (Veenman, Denessen, Van Den Akker, & Van Der Rijt, 2005), Veenman and his colleagues explored student cooperative behaviors in the dimensions of help seeking, help giving, and constructive activities. These behaviors were further examined in categories such as instrumental (e.g., requesting an explanation of process), executive (e.g., asking for a direct answer), confirmatory (e.g., verifying the proposed suggestion), and affective (e.g., giving positive comments on the collaboration process). The present study shows that cooperative learning can facilitate learners' peer collaboration and elaboration behaviors and that there are significant positive

relationships between these behaviors and student achievement. Future research is recommended to study what specific elaboration and peer collaboration behaviors in a cooperative learning setting have direct bearing on EFL learners' English language proficiency. The categories established by Veenman and his colleagues can be used as a guideline to create an observation protocol or questionnaire scales for future EFL research in this regard.

On top of motivation and use of learning strategies, grammar achievement can be examined from different perspectives as well. In this study grammar achievement was assessed with a paper-and-pencil test. Future studies should develop or find existing reliable and valid measurement tool to evaluate the effects of cooperative learning on students' grammar in oral communication. In addition, the grammar structures covered in the present study were at a relatively micro-level. Future studies should also examine the effects of cooperative learning on students' use of larger syntactic structures such as adverbial clauses in their writing.

Furthermore, new distinctions might be made in population characteristics in future studies. The present study has examined how the effects of cooperative learning depend on learners' prior English ability levels. It would also be worth exploring to examine whether the effects of cooperative learning on English grammar achievement, learning motivation, and learning strategy utilization depend on learners' personality and learning styles, for example, whether the learners are introverts or extraverts (Furnham & Heaven, 1999), imagers or verbalizers (Riding, Burton, Rees, & Sharratt, 1995; Riding & Wigley, 1997), field dependents or field independents (Crozier, 1997), activists, reflectors, theorists, or pragmatists (Honey & Mumford, 1992, 2000), or convergers,

divergers, assimilators, or accommodators (Kolb, 1984, 1999). In addition, it would be interesting to investigate whether the effects of cooperative learning on academic achievement depend on learners' motivation prior to the treatment. All these are additional dimensions that warrant exploring in future investigations.

When discussing the interaction effects of cooperative learning and prior English ability level earlier in this chapter, the researcher of the present study has postulated that the effects of cooperative learning depend on how students perceive their ranking levels in relation to their peers and subsequent social roles they choose to take as a result of the perceptions. To test the hypothesis, future studies are recommended using analytic techniques drawn from Critical Discourse Analysis (CDA) to analyze classroom and small group discourses. According to Fairclough (1995, 2003), CDA can be used to explore the relations between language and the social structures of those who use it through examination of speaking or written discourse. Via CDA techniques such as semantic roles, presuppositions, and modalities, the socially situated identities and role-taking of cooperative learners of different prior English ability levels can be investigated in depth, and a clearer understanding of what is occurring in the language learning events within the cooperative learning groups can be obtained (Gee, 2005; Lewis, 2001; Rogers, 2004; Rowe, 2004).

With regard to the length of study, a 12-week EFL program was implemented in the current research. Future longitudinal studies (e.g., one or two years) are warranted to examine the long-term effects of cooperative learning. Longitudinal studies are specially needed in order to gain a better understanding of the effect of cooperative learning on medium class ranking students. In the present study, cooperative learners in the middle

two quartiles received positive effects of cooperative learning on motivation and strategy use but not on grammar achievement. It is worth exploring to see if longer implementation of cooperative learning would ultimately facilitate enhancement of medium ranking learners' English grammar achievement. Additionally, it is recommended that data be collected at various points of the future longitudinal studies to allow examination into how the effects of cooperative learning and the length of an EFL program interplay.

Finally, in discussing the effects of cooperative learning on grammar achievement, the researcher of the current study has surmised that student anxiety in the cooperative group was probably lower than that in the whole-class group. To examine this postulation, further investigations on anxiety is necessary. Possible measurement tools include but not limit to the Foreign Language Classroom Anxiety Scales (FLCAS) (Horwitz, Horwitz, & Cope, 1986).

Conclusion

With English becoming a global language and the Taiwanese government putting much emphasis on improving English education and enhancing the citizens' English proficiency, the significance of learning English is widely recognized over the country. Nevertheless, English classes in Taiwan typically are established in a way with instructors as the center of the process and learning an individual business. Many students in Taiwan find instruction of this foreign language uninspiring, and many of them are really challenged to learn this foreign language. According to Shen (2002), Taiwanese students' scores on the Test of English as a Foreign Language (TOEFL) have declined

dramatically to the bottom of the international list. All these phenomena call for attention to seeking ways to improve EFL instructional pedagogy in Taiwan.

Although the findings of some recent research in second or foreign language education in several countries have indicated cooperative learning to be educationally significant for second language acquisition, robust studies on the effects of cooperative learning in Taiwan are still sparse, and cooperative learning still does not have the weight it warrants in the EFL curriculum in Taiwan's classrooms. The present study was an effort to contribute to the body of literature in this regard by examining the effects of cooperative learning on motivation, utilization of learning strategies, and grammar achievement. The findings of the study are summarized as follows:

1. The results of the study consistently indicate cooperative learning as more effective than whole-class instruction in promoting English grammar achievement, learning motivation, and learning strategy utilization.

2. The findings show a pattern that the effects of cooperative learning depend on learners' ranking in relation to peers rather than on their objective performing levels identified by measures such as standardized tests. It is possible that the effects rely on learners' perceptions of their relative ranking and consequent role-taking.

3. For subgroups of learners at different ranking levels in relation to peers, although cooperative learning facilitates motivation and learning strategy use for all subgroups, the effects are comparatively greater for the higher and lower ranking learners.

4. For subgroups of learners at different ranking levels in relation to peers, cooperative learning facilitates English grammar achievement gain of learners at the higher and lower ranking levels but not the middle ranking level.

5. When higher (i.e., "create" and evaluate") levels of cognitive activities are involved, learners who receive cooperative learning display higher English grammar achievement than those who receive whole-class instruction. When medium (i.e., "analyze" and "apply") and lower (i.e. "understand" and "remember") levels of cognitive activities are involved, cooperative learners and whole-class learners demonstrate similar English grammar achievement.

6. There are significant positive relationships among grammar achievement, motivation (including self-efficacy and task value), use of learning strategies (including elaboration and peer collaboration), and prior English ability level.

The cooperative learning model in the present study represents a fairly radical departure from many of the college EFL classroom instructional models currently in vogue in Taiwan. The present study contributes to existing literature not only by looking into the effects of cooperative learning on academic performance and motivational variables, but also by going further to examining students' out-of-class learning strategy application. Additionally, the present study extends findings of past investigations by zooming into the differential effects of cooperative learning for students of different class ranking levels. This study further expands the repertoire of knowledge on cooperative learning by exploring how this pedagogy influences EFL student achievement at different cognitive levels.

Many questions have been answered in the present study. Yet many more have emerged as the current study evolved and came to an end. Implications on theory, research, and practice have been presented and directions of future research proposed as a result of the study. The researcher asserts that the findings from the present study would provide EFL practitioners in Taiwan information and perspectives to improve curricula and instruction. The researcher also asserts that, with findings added to the existing literature in the field of cooperative learning in Taiwanese EFL instruction, other researchers can depart from the findings of the study and keep extending the repertoire of the research field.

REFERENCES

Abrami, P. C., Chambers, B., Lou, Y., Poulsen, C., & Spence, J. C. (1999). Within-class grouping: Evidence versus conjecture. *National Institute Economic Review, 169,* 105-108. Retrieved November 1, 2005, from http://www.questia.com

Abrami, P. C., Lou, Y., Chambers, B., Poulsen, C., & Spence, J. C. (2000). Why should we group students within-class for learning? *Educational Research and Evaluation, 6,* 158-179.

Adams, I. W. (2000). Exploring the efficacy of cooperative/collaborative learning: The experience of college ESL teachers (Doctoral dissertation, University of New Orleans, 2000). *Dissertation Abstracts International, 61* (04)*,* 1271A.

Adkins, D. C. (1974). *Test construction: Development and interpretation of achievement tests* (2nd ed.). Columbus, OH: Charles E. Merrill. Retrieved November 15, 2005, from http://www.questia.com

Ajzen, I. (1989). *Attitudes, personality, and behavior*. Chicago: Dorsey Press.

Anderson, L. W., & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching and assessing: A revision of bloom's taxonomy of educational objectives.* New York: Longman.

Arey, M.-J. (1993). French films: Pre-texts for teaching syntax. *Foreign Language Annals, 26,* 252-264.

Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The Jigsaw classroom*. Beverly Hills, CA: Sage.

Asher, J. (1977). Children learning another language: A developmental hypothesis. *Child Development, 58*, 24-32.

Au, K. H. (1993). *Literacy instruction in multicultural settings*. Fort Worth, TX: Harcourt Brace College.

Babcock, S. P. (1993). *The significance of cultural influence within ESL/EFL classroom: A Taiwan experience.* Paper presented at the International Conference on Teacher Education in Second Language Teaching, Hong Kong. (ERIC Document Reproduction Service No. ED375681)

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Barnitz, J. G. (1998). Revising grammar instruction for authentic composing and comprehending. *The Reading Teacher, 51,* 608-611.

Bejarano, Y. (1987). Cooperative small-group methodology in the language classroom. *TESOL Quarterly, 21,* 483-501.

Bernard, L. L. (1926). *An introduction to social psychology*. New York: Henry Holt and Company.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28,* 69-83.

Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall.

Bruffee, K. A. (1999). *Collaborative learning: Higher education, interdependence, and the authority of knowledge* (2nd ed.). Baltimore: Johns Hopkins University Press.

Bruner, J. (1960). *The process of education*. Cambridge, MA: Harvard University Press.

Bruner, J. (1990). *Act of meaning*. Cambridge, MA: Harvard University Press.

Calderon, M., Hertz-Lazarowitz, R., & Slavin, R. (1998). Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *Elementary School Journal, 99,* 153-165.

Calkins, L. M. (1986). *The art of teaching writing*. Portsmouth, NH: Heinemann.

Cambridge advanced learner's dictionary (2003). Cambridge, United Kingdom: Cambridge University Press.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1,* 1-47.

Chang, K.-Y. R., & Smith, Wm. F. (1991). Cooperative learning and CALL/IVD in beginning Spanish: An experiment. *The Modern Language Journal, 75,* 205-211.

Chang, M.-M. (1995). Teacher-oriented learning vs. cooperative learning in English reading class. *Bulletin of National Pingtung Polytechnic Institute, 4,* 271-277.

Chen, H.-C. (1998). The performance of junior college students studying English through cooperative learning. *Proceedings of the Seventh International Symposium on English Teaching* (pp. 231-240). Taipei, Taiwan: Crane.

Chen, H.-C. (1999). A comparison between cooperative learning and traditional, whole-class methods—Teaching English in a junior college. *Academic Journal of Kang-Ning, 3,* 69-82.

Cheng, Y.-S. (1998). A qualitative inquiry of second language anxiety: Interviews with Taiwanese EFL students. *The Proceedings of the Seventh International Symposium on English Teaching* (pp. 309-320). Taipei, Taiwan: Crane.

Chiu, A. (Ed.) (2000). *Longman dictionary of contemporary English* (5th ed.). Hong Kong: Longman.

Chu, M. P. (1996, March). An application of the Jigsaw method in a freshman English class. *Hwang-Kung Foreign Languages Journal,* 207-222.

Clarke, J. (1999). Pieces of the puzzle: The Jigsaw method. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 34-50). Westport, CT: Praeger. Retrieved October 31, 2005, from http://www.questia.com

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved October 22, 2005, from http://www.questia.com

Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge, United Kingdom: Cambridge University Press.

Covington, M. V., & Omelich, C. L. (1979). Effort: the double-edged sword in school achievement. *Journal of Educational Psychology, 71,* 169-182.

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Prentice Hall.

Crookes, G. (1991). Second language speech production research: A methodologically oriented review. *Studies in Second Language Acquisition, 13,* 113-132.

Crozier, W. R. (1997). *Individual learners: Personality differences in education*. London: Routledge.

Cudd, E. T., & Roberts, L. L. (1994). A scaffolding technique to develop sentence sense and vocabulary. *The Reading Teacher, 47,* 346-349.

Dansereau, D. F. (1988). Cooperative learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 103-120). Orlando, FL: Academic Press.

De Graaff, R. (1997a). *Differential effects of explicit instruction on second language acquisition*. Leiden, Netherlands: Holland Institute of Generative Linguistics.

De Graaff, R. (1997b). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19,* 249-276.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26,* 325-346.

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations, 2,* 129-152.

Dewey, J. (1916). *Democracy and education*. New York: Macmillan.

Dewey, J. (1924). *Democracy and education*. New York: Macmillan.

Dewey, J. (1963). *Experience and education*. New York: Collier Books.

Dewey, J. (1964). My pedagogic creed. In L. A. Fiedler & J. Vinocur (Eds.), *The continuing debate: Essays on education* (pp. 169-181). New York: St Martin's Press.

Diggins, P. B. (1997). *The European dimension in the curriculum of the primary school and the junior certificate.* Retrieved May 10, 2005, from http://www.eduvinet.de/eduvinet/irl007.htm

Dornyei, Z. (2001). *Teaching and researching motivation.* Harlow, England: Longman.

Duckworth, E. (1964). Piaget rediscovered. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered: A report of the conference on cognitive studies and curriculum development* (pp. 1-5). Ithaca, NY: Cornell University.

Earley, P. C., & Kanfer, R. (1985). The influence of component participation and role models on goal acceptance, goal satisfaction and performance. *Organizational Behavior and Human Performance, 36,* 378-390.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11,* 135-172.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21,* 215-225.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109-132.

Eccles (Parsons), J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). San Francisco: W. H. Freeman.

Educational Testing Service. (2005). *TOEFL—Test of English as a foreign language.* Retrieved October 27, 2005, from ETS Web site: http://www.ets.org

Ellis, R. (1993a). Interpretation-based grammar teaching. *System, 21,* 69-78.

Ellis, R. (1993b). The structural syllabus and second language acquisition. *TESOL Quarterly, 27,* 91-113.

Enginarlar, H. (1994). Sentence combining plus: A new use for an old technique. *ELT Journal, 48,* 214-224.

Erez, M. (1977). Feedback: A necessary condition for the goal setting-performance relationship. *Journal of Applied Psychology, 62,* 624-627.

Evans, R., Venetozzi, R., Bundrick, M., & McWilliams, E. (1988). The effects of sentence-combining instructions on writing and on standardized test scores. *Journal of Educational Research, 82,* 53-57.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. New York: Longman.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. London: Routledge. Retrieved November 21, 2005, from http://www.questia.com

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fotos, S., & Ellis, R. (1991). Communication about grammar: A task-based approach. *TESOL Quarterly, 25,* 605-628.

Fuchs, M. (2002). *Grammar express teacher's manual*. White Plains, NY: Longman.

Fuchs, M., & Bonner, M. (2001). *Grammar express*. White Plains, NY: Longman.

Furnham, A., & Heaven, P. (1999). *Personality and social behaviour*. London: Arnold.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman.

Gall, M. D., Borg, W. R., & Gall, J. P. (2003). *Educational research: An introduction* (7th ed.). White Plains, NY: Allyn and Bacon.

Garland, H., & Adkinson, J. H. (1987). Standards, persuasion, and performance. *Group and Organization Studies, 12,* 208-220.

Gass, S. M., & Selinker, L. (1994). *Second language acquisition: An introductory course*. Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved December 6, 2005, from http://www.questia.com

Gee, J. P. (2005). *An introduction to discourse analysis: Theory and method* (2nd ed.). New York: Routledge. Retrieved November 22, 2005, from http://www.questia.com

Ghaith, G. M. (2001). Learners' perceptions of their STAD cooperative experience. *System, 29,* 289-301.

Ghaith, G. M. (2002). The relationship between cooperative learning, perception of social support, and academic achievement. *System, 30,* 263-273.

Ghaith, G. M. (2003a). 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,* 451-474.

Ghaith, G. M. (2003b). The relationship between forms of instruction, achievement and perception of classroom climate. *Educational Research, 45,* 83-93.

Ghaith, G. M. (2004). Correlates of the implementation of the STAD cooperative learning method in the English as a foreign language classroom. *International Journal of Bilingual Education and Bilingualism, 7,* 279-294.

Ghaith, G. M., & Bouzeineddine, A. R. (2003). Relationship between reading attitudes, achievement, and learners' perceptions of their Jigsaw II cooperative learning experience. *Reading Psychology, 24,* 105-121.

Ghaith, G. M., & El-Malak, M. A. (2004). Effect of Jigsaw II on literal and higher order EFL reading comprehension. *Educational Research and Evaluation, 10,* 105-115.

Ghaith, G. M., & Yaghi, H. M. (1998). Effect of cooperative learning on the acquisition of second language rules and mechanics. *System, 26,* 223-234.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston: Allyn and Bacon.

Glesne, C. (1999). *Becoming qualitative researchers: An introduction* (2nd ed.). New York: Longman.

Government Information Office. (2003). *Taiwan yearbook 2003.* Retrieved June 26, 2004, from GIO Web site: http://www.gio.gov.tw/taiwan-website/5-gp/yearbook

Government Information Office. (2004). *Taiwan yearbook 2004.* Retrieved October 15, 2005, from GIO Web site: http://www.gio.gov.tw/taiwan-website/5-gp/yearbook

Government Information Office. (2005). *Taiwan at a glance 2005-2006.* Retrieved December 5, 2005, from GIO Web site: http://www.gio.gov.tw/taiwan-website/5-gp/glance

Gravetter, F. J., & Wallnau, L. B. (2004). *Statistics for the behavioral sciences* (6th ed.). Belmont, CA: Wadsworth.

Green, S. B., Salkind, N. J., & Akey, T. M. (2000). *Using SPSS for windows: Analyzing and understanding data* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Gunderson, B., & Johnson, D. (1980). Building positive attitudes by using cooperative learning groups. *Foreign Language Annals, 13,* 39-43.

Hammerley, H. (1987). The immersion approach: Litmus test of second language acquisition through classroom communication. *The Modern Language Journal, 71,* 395-401.

Hammond, R. (1988). Accuracy versus communicative competency: The acquisition of grammar in the second language classroom. *Hispania, 71*, 408-417.

Hansman, C. A. (2001). Context-based adult learning. In S. B. Merriam (Ed.), *The new update on adult learning theory* (pp. 43-51). San Francisco: Jossey-Bass.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Higgs, T. V., & Clifford, R. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57-79). Skokie, IL: National Textbook.

Hinkle, D., Wiersma, W., & Jurs, S. G. (1994). *Applied statistics for the behavioral sciences* (3rd ed.). Boston: Houghton Mifflin.

Ho, Y.-K. (2003). Audiotaped dialogue journals: An alternative form of speaking practice. *ELT Journal, 57,* 269-277.

Holt, D. D. (Ed.). (1993). *Cooperative learning: A response to linguistic and cultural diversity*. Washington, DC: Center for Applied Linguistics.

Honey, P., & Mumford, A. (1992). *The manual of learning styles* (3rd ed.). Maidenhead, United Kingdom: Peter Honey.

Honey, P., & Mumford, A. (2000). *The learning styles helper's guide*. Maidenhead, United Kingdom: Peter Honey.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70,* 125-132.

Huck, S. W. (2004). *Reading statistics and research* (4th ed.). Boston: Allyn and Bacon.

Isik, A. (2000). The role of input in second language acquisition: More comprehensible input supported by grammar instruction or more grammar instruction? *I.T.L. Review of Applied Linguistics, 129-130,* 225-274.

Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research, 50,* 438-460.

Johnson, D. W., & Johnson, F. P. (2003). *Joining together: Group theory and group skills* (8th ed.). Needham Heights, MA: Allyn and Bacon.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book.

Johnson, D. W., & Johnson, R. T. (1990). Social skills for successful group work. *Educational Leadership, 47,* 29-33.

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., & Roy, P. (1984). *Circles of learning: Cooperation in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). *Cooperative learning : Increasing college faculty instructional productivity* (ASHE-ERIC Higher Education Report No. 4.) Washington, DC: School of Education and Human Development, George Washington University.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: What evidences is there that it works? *Change, 30,* 26-35.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (1989). Impact of goal and resource interdependence on problem-solving success. *Journal of Social Psychology, 129,* 621-629.

Johnson, D. W., Johnson, R. T., Stanne, M. B., & Garibaldi, A. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology, 130,* 507-516.

Kagan, S. (1985). Co-op Co-op: A flexible cooperative learning technique. In R. E. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 437-452). New York: Plenum Press.

Kerr, N., & Bruun, S. (1983). The dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology, 44,* 78-94.

Kessler, C. (Ed.). (1992). *Cooperative language learning: A teacher's resource book*. Englewood Cliffs, NJ: Prentice Hall.

Kitchener, R. F. (1991). Jean Piaget: the unknown sociologist? *BJS, 42,* 421-442.

Kohn, A. (1991a). Don't spoil the promise of cooperative learning. *Educational Leadership, 48,* 93-94.

Kohn, A. (1991b). Group grade grubbing versus cooperative learning. *Educational Leadership, 48,* 83-87.

Kolb, D. A. (1984). *Experimental learning*. Englewood Cliffs, NJ: Prentice-Hall.

Kolb, D. A. (1999). *Kolb learning style inventory* (3rd ed.). Philadelphia: Hay Group.

Krashen, S. D. (1976). Formal and informal linguistic environments in language acquisition and language learning. *TESOL Quarterly, 10,* 157-168.

Krashen, S. D. (1982). *Principles and practices of second language acquisition*. Oxford, United Kingdom: Pergamon Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. D. (1988). *Second language acquisition and second language learning*. Hertfordshire, United Kingdom: Prentice Hall.

Krashen, S. D. (2002). Theory versus practice in language training. In V. Zamel & R. Spack (Eds.), *Enriching ESOL pedagogy: Reading and activities for engagement, reflection, and inquiry* (pp. 211-228). Mahwah, NJ: Lawrence Erlbaum Associates.

Krashen, S. D. (2003). *Explorations in language acquisition and use*. Portsmouth, NH: Heinemann.

Krashen, S. D., Butler, J., Birnbaum, R., & Robertson, J. (1978). Two studies in language acquisition and language learning. *I.T.L. Review of Applied Linguistics, 39-40,* 73-79.

Lai, S.-H. (2001). Student perceptions regarding needs for communication-based English instruction in a selected university in southern Taiwan (Doctoral Dissertation, University of South Dakota, 2001). *Dissertation Abstracts International, 62* (03), 900A.

Language Training and Testing Center. (n.d.). *A brief introduction to the GEPT.* Retrieved January 27, 2005, from http://www.lttc.ntu.edu.tw/E_LTTC/gept_eng_main.htm

Lantolf, J. P., & Appel, G. (Eds.). (1994). *Vygotskian approaches to second language research*. Westport, CT: Ablex.

Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston: Heinle.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Use and interpretation* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved October 15, 2005, from http://www.questia.com

Levine, G. (1991). *A guide to SPSS for analysis of variance*. Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved October 23, 2005, from http://www.questia.com

Lewis, C. (2001). *Literacy practices as social acts: Power, status and cultural norms in the classroom*. Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved November 22, 2005, from http://www.questia.com

Liang, T. (2002). *Implementing cooperative learning in EFL teaching: Process and effects*. Doctoral Dissertation, National Taiwan Normal University, Taipei, Taiwan.

Liao, H.-C., & Chang, F. (2003). Writer community in an English composition class at college level in Taiwan: Implementation of collaborative learning. In *Proceedings of the 2003 National Conference on English Teaching and Learning.* Yunlin, Taiwan: National Huwei Institute of Technology.

Liao, H.-C., & Hsueh, Y.-U. (2005). *Collaborative groups in Taiwan's ESL writing classes.* Paper presented at the 39th Annual Teaching English to Speakers of Other Languages Convention and Exhibit (TESOL 2005) Graduate Student Forum, San Antonio, TX.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12,* 429-448.

Lo, Y. H. (1998). 合作學習: 談二專英文的教材與教法 [Cooperative learning: Materials and methods for a two-year junior college English class]. In *Proceedings of the seventh international symposium on English teaching* (pp. 727-743). Taipei, Taiwan: English Teachers' Association and Crane.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Lomax, R. G. (2001a). *An introduction to statistical concepts for education and behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved October 17, 2005, from http://www.questia.com

Lomax, R. G. (2001b). *Statistical concepts: A second course for education and the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved October 15, 2005, from http://www.questia.com

Long, M. H. (1983). Does second language instruction make a difference? A review of research. *TESOL Quarterly, 17,* 359-382.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

Lou, Y., Abrami, P. C., & Spence, John C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *Journal of Educational Research, 94,* 101-112. Retrieved November 1, 2005, from http://www.questia.com

Maller, J. B. (1929). *Cooperation and competition: An experimental study in motivation.* Unpublished doctoral dissertation, Columbia University, New York.

Matsui, T., Kakuyama, T., & Onglatco, M. L. (1987). Effects of goals and feedback on performance in groups. *Journal of Applied Psychology, 72,* 407-415.

McLaughlin, B. (1987). *Theories of second-language learning*. London: Edward Arnold.

McLaughlin, B. (1990a). "Conscious" versus "unconscious" learning. *TESOL Quarterly, 24,* 617-634.

McLaughlin, B. (1990b). Restructuring. *Applied Linguistics, 11,* 113-128.

McMillan, J. H., & Schumacher, S. (1997). *Research in education: A Conceptual introduction* (4th ed.). New York: Longman.

Merriam, S. B., & Caffarella, R. S. (1999). *Learning in adulthood: A comprehensive guide* (2nd ed.). San Francisco: Jossey-Bass.

Milleret, M. (1992). Cooperative learning in the Portuguese-for-Spanish-speakers classroom. *Foreign Language Annals, 25,* 435-440.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2002). On the roles of task model variables in assessment design. In S. H. Irvine & P. C. Kyllonen (Eds.), *Item generation for test development* (pp. 97-128). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved November 15, 2005, from http://www.questia.com

Mitchell, M. L., & Jolley, J. M. (2004). *Research design explained* (5th ed.). Belmont, CA: Wadsworth.

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11,* 247-266.

Mueller, M. E. (1983). *The effects of goal setting and competition on performance: A laboratory study*. Unpublished master's thesis, University of Minnesota, Minneapolis.

Nassaji, H. (2000). Toward integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Modern Language Journal, 84,* 241-250.

Neville, D. D., & Searls, E. F. (1991). A meta-analytical review of the effect of sentence-combining on reading comprehension. *Reading Research and Instruction, 31,* 63-76.

Nicola, M. (1990). Experimenting with the new methods. *Dialog on Language Instruction, 6*, 61-72.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50,* 417-528.

Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning, 51,* 157-213.

Nunan, D. (1999). *Second language teaching and learning*. Boston: Heinle & Heinle.

O'Hare, F. (1973). *Sentence combining: Improving student writing without formal grammar instruction*. Urbana, IL: National Council of Teachers of English.

Olsen, R. E., & Kagan, S. (1992). About cooperative learning. In C. Kessler (Ed.), *Cooperative language learning: A teacher's resource book* (pp. 1-30). Englewood Cliffs, NJ: Prentice Hall.

Orem, R. A. (2001). Journal writing in adult ESL: Improving practice through reflective writing. *New Directions for Adult and Continuing Education, 90,* 69-77.

Oxford, R. L. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal, 81,* 443-456.

Pappas, T. (2003). Encyclopedia Britannica 2003 Ready Reference (Version 2003.0.0.0) [Computer software]. Chicago: Encyclopedia Britannica.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace College.

Pepitone, E. A. (1980). *Children in cooperation and competition: Toward a developmental social psychology*. Lexington, MA: Lexington Books.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.

Peyton, J. K., & Reed, L. (1990). *Dialogue journal writing with nonnative English speakers: A handbook for teachers*. Alexandria, VA: Teachers of English to Speakers of Other Languages.

Piaget, J. (1932). *The moral judgment of the child* (M. Gabain, Trans.). New York: Harcourt, Brace and Company.

Piaget, J. (1950). *The psychology of intelligence*. London: Routledge and Paul.

Piaget, J. (1964). Development and learning. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered: A report of the conference on cognitive studies and curriculum development* (pp. 7-20). Ithaca, NY: Cornell University.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44,* 493-527.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16,* 385-407.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Rep. No. 91-B-004). Ann Arbor, MI: University of Michigan.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement, 53*, 801-813.

Rakestraw, T. L., & Weiss, H. M. (1981). The interaction of social influences and task experience on goals, performance, and performance satisfaction. *Organizational Behavior and Human Performance, 27,* 326-344.

Riding, R., Burton, D., Rees, G., & Sharratt, M. (1995). Cognitive style and personality in 12-year-old children. *British Journal of Educational Psychology, 65,* 113-124.

Riding, R., & Wigley, S. (1997). The relationship between cognitive style and personality in further education students. *Personality and Individual Differences, 23,* 379-389.

Robinson, T. A. (1995). *Aristotle in outline*. Indianapolis, IN: Hackett.

Rogers, R. (2004). A critical discourse analysis of literate identities across contexts: Alignment and conflicts. In R. Rogers (Ed.), *An introduction to critical discourse analysis in education* (pp. 51-78). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved November 23, 2005, from http://www.questia.com

Rowe, S. (2004). Discourse in activity and activity as discourse. In R. Rogers (Ed.), *An introduction to critical discourse analysis in education* (pp. 79-96). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved November 23, 2005, from http://www.questia.com

Rylant, C. (1982). *When I was young in the mountains*. New York: Dutton.

Savignon, S. J. (1972). *Communicative competence: An experiment in foreign language teaching*. Philadelphia: Center for Curriculum Development.

Savignon, S. J. (1991). Communicative language teaching: State of art. *TESOL Quarterly, 25,* 261-277.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129-158.

Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13,* 206-226.

Schmidt, R. W. (1995). *Attention and awareness in foreign language learning*. Honolulu: Second Language Teaching and Curriculum Center, University of Hawaii at Manoa.

Schulz, R. (1991). Second language acquisition theories and teaching practice: How do they fit? *The Modern Language Journal, 75,* 17-26.

Sellers, J. A. (2005). Using cooperative learning in a content-based Spanish course: The Latin American *telenovela* (Doctoral dissertation, University of Wyoming, 2005). *Dissertation Abstracts International, 66* (03), 864A.

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research, 50,* 241-271.

Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics, 11,* 159-168.

Shen, H.-Y. (2002, November 8). Student English writing extension center established in Taiwan. In *Liberty Times*. Retrieved October 27, 2005, from http://www.libertytimes.com.tw/2002/new/nov/8/today-il.htm

Slavin, R. E. (1991a). Group rewards make groupwork work. *Educational Leadership, 48,* 89-91.

Slavin, R. E. (1991b). Synthesis of research on cooperative learning. *Educational Leadership, 48,* 71-82.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Needham Heights, MS: Allyn and Bacon.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21,* 43-69.

Snowman, J., & Biehler, R. (2005). *Psychology applied to teaching* (11th ed.). Boston: Houghton Mifflin.

Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics, 6,* 239-254.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15,* 205-224.

Speaker, R. B., Jr., & Speaker, P. R. (1991). Sentence collecting: Authentic literacy events in the classroom. *Journal of Reading, 35,* 92-95.

Staton, J., Shuy, R. W., Peyton, J. K., & Reed, L. (1988). *Dialogue journal communication: Classroom, linguistic, social, and cognitive views*. Norwood, N.J.: Ablex.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved October 29, 2005, from http://www.questia.com

Stevens, J. P. (1999). *Intermediate statistics: A modern approach* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved October 17, 2005, from http://www.questia.com

Su, S. F. (2003). *A study on English instructional innovation in Taiwan junior high schools—Exploring the status quo and feasible coping strategies.* Doctoral Dissertation, National Taiwan Normal University, Taipei, Taiwan.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-124). Oxford, United Kingdom: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkle (Ed.), *Handbook on research in second language teaching and learning* (pp. 471-484). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82,* 320-337.

Swezey, R. W., Meltzer, A. L., & Salas, E. (1994). Some issues involved in motivating teams. In H. F. O'Neil, Jr. & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 141-169). Hillsdale, NJ: Lawrence Erlbaum Associates.

Szostek, C. (1994). Assessing the effects of cooperative learning in an honors foreign language classroom. *Foreign Language Annals, 27,* 252-261.

Tomasello, M., & Herron, C. (1989). Feedback for language transfer errors: The Garden Path technique. *Studies in Second Language Acquisition, 11,* 385-395.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16,* 183-203.

Triplett, N. (1897). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology, 9,* 507-533.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 271-360). New York: Academic Press.

Van Baxtel, C. (2000). *Collaborative concept learning: Collaborative learning tasks, student interaction, and the learning of physics concepts.* Unpublished doctoral dissertation, Rtrecht University, Rtrecht, Netherlands.

Vaughan, W. (2002). Effects of cooperative learning on achievement and attitude among students of color. *Journal of Educational Research, 95,* 359-364.

Veenman, S., Denessen, E., Van Den Akker, A., & Van Der Rijt, J. (2005). Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. *American Educational Research Journal, 42*, 115-151.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walling, D. R. (1987). *A model for teaching writing: Process and product*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Wang, Y.-C. H. (2001). Using cooperative learning in English conversation course for junior college students in Taiwan. *Journal of Penghu Institute of Technology, 4,* 333-356.

Weaver, C. (1990). *Understanding whole language: From principles to practice*. Portsmouth, NH: Heinemann.

Weaver, C. (1996). *Teaching grammar in context*. Portsmouth, NH: Heinemann.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31,* 369-395.

Webb, N. M., & Farivar, S. (1999). Developing productive group interaction in middle school mathematics. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 117-149). Mahwah, NJ: Lawrence Erlbaum Associates.

Webb, N. M., Farivar, S. H., & Mastergeorge, A. M. (2002). Productive helping in cooperative groups. *Theory into practice, 41,* 13-20.

Webb, N. M., & Mastergeorge, A. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39,* 73-97.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York: Simon and Schuster.

Wei, C.-L. (1996). 國內大專學生對「合作學習」英語教學活動的看法 [Cooperative English learning activities: Perceptions of Taiwanese college students]. *教育研究資訊 [Educational Research], 4,* 13-26.

Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review, 12,* 1-14.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics, 8,* 95-110.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*, 49-78.

Yager, S., Johnson, R. T., Johnson, D. W., & Snider, B. (1986). Impact of group processing on achievement in cooperative learning groups. *Journal of Social Psychology, 126,* 389-397.

Yu, G. K. (1993). Cooperative learning for heterogeneous EFL classrooms in secondary schools in Taiwan. *Journal of Kaohsiung Normal University, 4*, 209-229.

APPENDIXES

Appendix A

Confirmatory Factor Analyses with the Lambda-ksi Estimates on the Motivation Items of the MSLQ

Items	Scale*	LX estimate
In a class like this, I prefer course material that really challenges me so I can learn new things.	1	.64
In a class like this, I prefer course material that arouses my curiosity, even if it is difficult to learn.	1	.69
The most satisfying thing for me in this course is trying to understand the content as thoroughly as possible.	1	.66
When I have the opportunity in this class, I choose course assignments that I can learn from even if they don't guarantee a good grade.	1	.55
Getting a good grade in this class is the most satisfying thing for me right now.	2	.71
The most important thing for me right now is improving my overall grade point average, so my main concern in this class is getting a good grade.	2	.58
If I can, I want to get better grades in this class than most of the other students.	2	.48
I want to do well in this class because it is important to show my ability to my family, friends, employer, or others.	2	.44
I think I will be able to use what I learn in this course in other courses.	3	.57
It is important for me to learn the course material in this class.	3	.64
I am very interested in the content area of this course	3	.88
I think the course material in this class is useful for me to learn.	3	.86
I like the subject matter of this course.	3	.88
Understanding the subject matter of this course is very important to me.	3	.84
If I study in appropriate ways, then I will be able to learn the material in this course.	4	.57
It is my own fault if I don't learn the material in this course.	4	.38
If I try hard enough, then I will understand the course material.	4	.84
If I don't understand the course material, it is because I didn't try hard enough.	4	.47
I believe I will receive an excellent grade in this class.	5	.83

(Table Continued)

Items	Scale*	LX estimate
I am certain I can understand the most difficult material presented in the readings for this course.	5	.70
I am confident I can understand the basic concepts taught in this course.	5	.63
I am confident I can understand the most complex material presented by the instructor in this course.	5	.71
I am confident I can do an excellent job on the assignments and tests in this course.	5	.86
I expect to do well in this class.	5	.89
I am certain I can master the skills being taught in this class.	5	.77
Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.	5	.87
When I take a test I think about how poorly I am doing compared with other students.	6	.60
When I take a test I think about items on other parts of the test I can't answer.	6	.42
When I take tests I think of the consequences of failing.	6	.62
I have an uneasy, upset feeling when I take an exam.	6	.88
I feel my heart beating fast when I take an exam.	6	.76

Note. Scale 1: Intrinsic Goal Orientation, Scale 2: Extrinsic Goal Orientation, Scale 3: Task Value, Scale 4: Control of Learning Beliefs, Scale 5: Self-Efficacy for Learning and Performance, Scale 6: Test Anxiety. N = 356.

Appendix B

Confirmatory Factor Analyses with the Lambda-ksi Estimates on the Learning Strategies Items of the MSLQ

Item	Scale*	LX estimate
When I study for this class, I practice saying the material to myself over and over.	7	.62
When studying for this class, I read my class notes and the course reading over and over again.	7	.63
I memorize key words to remind me of important concepts in this class.	7	.56
I make lists of important terms for this course and memorize the lists.	7	.58
When I study for this class, I pull together information from different sources, such as lectures, readings, and discussions.	8	.60
I try to relate ideas in this subject to those in other courses whenever possible.	8	.60
When reading for this class, I try to relate the material to what I already know.	8	.74
When I study for this course, I write brief summaries of the main ideas from the readings and the concepts from the lectures.	8	.42
I try to understand the material in this class by making connections between the readings and the concepts from the lectures.	8	.71
I try to apply ideas from course readings in other class activities such as lecture and discussion.	8	.65
When I study the readings for this course, I outline the material to help me organize my thoughts.	9	.57
When I study for this course, I go through the readings and my class notes and try to find the most important ideas.	9	.55
I make simple charts, diagrams, or tables to help me organize course material.	9	.45
When I study for this course, I go over my class notes and make an outline of important concepts.	9	.75
I often find myself questioning things I hear or read in this course to decide if I find them convincing.	10	.49
When a theory, interpretation, or conclusion is presented in class or in the readings, I try to decide if there is good supporting evidence.	10	.76
I treat the course material as a starting point and try to develop my own ideas about it.	10	.66

(Table Continued)

Item	Scale*	LX estimate
I try to play around with ideas of my own related to what I am learning in this course.	10	.74
Whenever I read or hear an assertion or conclusion in this class, I think about possible alternatives.	10	.67
During class time I often miss important points because I am thinking of other things. (REVERSED)	11	.40
When reading for this course, I make up questions to help focus my reading.	11	.44
When I become confused about something I am reading for this class, I go back and try to figure it out.	11	.47
If course materials are difficult to understand, I change the way I read the material.	11	.54
Before I study new course material thoroughly, I often skim it to see how it is organized.	11	.53
I ask myself questions to make sure I understand the material I have been studying in this class.	11	.58
I try to change the way I study in order to fit the course requirements and instructor's teaching style.	11	.43
I often find that I have been reading for class but don't know what it was all about. (REVERSED)	11	.35
I try to think through a topic and decide what I am supposed to learn from it rather than just reading it over when studying.	11	.60
When studying for this course I try to determine which concepts I don't understand well.	11	.61
When I study for this class, I set goals for myself in order to direct my activities in each study period.	11	.55
If I get confused taking notes in class, I make sure I sort it out afterwards.	11	.50
I usually study in a place where I can concentrate on my course work.	12	.52
I make good use of my study time for this course.	12	.81
I find it hard to stick to a study schedule. (REVERSED)	12	.52
I have a regular place set aside for studying.	12	.56

(Table Continued)

Item	Scale*	LX estimate
I make sure I keep up with the weekly reading and assignments for this course.	12	.64
I attend class regularly.	12	.37
I often find that I don't spend very much time on this course because of other activities. (REVERSED)	12	.48
I rarely find time to review my notes or readings before an exam. (REVERSED)	12	.40
I often feel so lazy or bored when I study for this class that I quit before I finish what I planned to do. (REVERSED)	13	.53
I work hard to do well in this class even if I don't like what we are doing.	13	.65
When course work is difficult, I give up or only study the easy parts. (REVERSED)	13	.52
Even when course materials are dull and uninteresting, I manage to keep working until I finish.	13	.74
When studying for this course, I often try to explain the material to a classmate or a friend.	14	.54
I try to work with other students from this class to complete the course assignment.	14	.82
When studying for this course, I often set aside time to discuss the course material with a group of students from the class.	14	.84
Even if I have trouble learning the material in this class, I try to do the work on my own, without help from anyone. (REVERSED)	15	.20
I ask the instructor to clarify concepts I don't understand well.	15	.17
When I can't understand the material in this course, I ask another student in this class for help.	15	.90
I try to identify students in this class whom I can ask for help if necessary.	15	.79

Note. Scale 7: Rehearsal, Scale 8: Elaboration, Scale 9: Organization, Scale 10: Critical Thinking, Scale 11: Metacognitive Self-Regulation, Scale 12: Time and Study Environment, Scale 13: Effort Regulation, Scale 14: Peer Learning, Scale 15: Help Seeking. N = 356.

Appendix C

MSLQ Pretest Questionnaire

INSTRUCTION: The following questions ask about your motivation and learning strategies. Remember there are no right or wrong answers; just answer as accurately as possible. If you think the statement is always true of you, circle 7; if a statement is never true of you, circle 1. If the statement is more or less true of you, find the number between 1 and 7 that best describes you. The survey is for research purpose only. Your course instructor will <u>not</u> have access to your answers, and it will <u>not</u> affect your course grade in any way.	1 Never True of Me 2 Rarely True of Me 3 Occasionally True of Me 4 Sometimes True of Me 5 Often True of Me 6 Usually True of Me 7 Always True of Me
Part I. Please answer items 1-14 based on how you <u>feel about this course now</u>.	
1. I think I will be able to use what I learn in this course in other courses.	1 2 3 4 5 6 7
2. It is important for me to learn the course material in this class.	1 2 3 4 5 6 7
3. I am very interested in the content area of this course.	1 2 3 4 5 6 7
4. I think the course material in this class is useful for me to learn.	1 2 3 4 5 6 7
5. I like the subject matter of this course.	1 2 3 4 5 6 7
6. Understanding the subject matter of this course is very important to me.	1 2 3 4 5 6 7
7. I believe I will receive an excellent grade in this class.	1 2 3 4 5 6 7
8. I am certain I can understand the most difficult material presented in the readings for this course.	1 2 3 4 5 6 7
9. I am confident I can understand the basic concepts taught in this course.	1 2 3 4 5 6 7
10. I am confident I can understand the most complex material presented by the instructor in this course.	1 2 3 4 5 6 7
11. I am confident I can do an excellent job on the assignments and tests in this course.	1 2 3 4 5 6 7
12. I expect to do well in this class.	1 2 3 4 5 6 7
13. I am certain I can master the skills being taught in this class.	1 2 3 4 5 6 7
14. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.	1 2 3 4 5 6 7

Part II. Please answer items 15-23 based on your general study habits.	
15. When I study for a class, I pull together information from different sources, such as lectures, readings, and discussions.	1 2 3 4 5 6 7
16. I try to relate ideas in the subject of a course to those in other courses whenever possible.	1 2 3 4 5 6 7
17. When reading for a class, I try to relate the material to what I already know.	1 2 3 4 5 6 7
18. When I study for a course, I write brief summaries of the main ideas from the readings and the concepts from the lectures.	1 2 3 4 5 6 7
19. I try to understand the material in a class by making connections between the readings and the concepts from the lectures.	1 2 3 4 5 6 7
20. I try to apply ideas from course readings in other class activities such as lecture and discussion.	1 2 3 4 5 6 7
21. When studying for a course, I often try to explain the material to a classmate or a friend.	1 2 3 4 5 6 7
22. I try to work with other students from the same class to complete the course assignment.	1 2 3 4 5 6 7
23. When studying for a course, I often set aside time to discuss the course material with a group of students from the class.	1 2 3 4 5 6 7

Appendix D

MSLQ Posttest Questionnaire

INSTRUCTION: The following questions ask about your motivation and learning strategies. Remember there are no right or wrong answers; just answer as accurately as possible. If you think the statement is always true of you, circle 7; if a statement is never true of you, circle 1. If the statement is more or less true of you, find the number between 1 and 7 that best describes you. The survey is for research purpose only. Your course instructor will <u>not</u> have access to your answers, and it will <u>not</u> affect your course grade in any way.	Never True of Me Rarely True of Me Occasionally True of Me Sometimes True of Me Often True of Me Usually True of Me Always True of Me
Part I. Please answer items 1-14 based on how you <u>feel about this course now</u>.	
1. I think I will be able to use what I learn in this course in other courses.	1 2 3 4 5 6 7
2. It is important for me to learn the course material in this class.	1 2 3 4 5 6 7
3. I am very interested in the content area of this course.	1 2 3 4 5 6 7
4. I think the course material in this class is useful for me to learn.	1 2 3 4 5 6 7
5. I like the subject matter of this course.	1 2 3 4 5 6 7
6. Understanding the subject matter of this course is very important to me.	1 2 3 4 5 6 7
7. I believe I will receive an excellent grade in this class.	1 2 3 4 5 6 7
8. I am certain I can understand the most difficult material presented in the readings for this course.	1 2 3 4 5 6 7
9. I am confident I can understand the basic concepts taught in this course.	1 2 3 4 5 6 7
10. I am confident I can understand the most complex material presented by the instructor in this course.	1 2 3 4 5 6 7
11. I am confident I can do an excellent job on the assignments and tests in this course.	1 2 3 4 5 6 7
12. I expect to do well in this class.	1 2 3 4 5 6 7
13. I am certain I can master the skills being taught in this class.	1 2 3 4 5 6 7
14. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.	1 2 3 4 5 6 7

Part II. Please answer items 15-23 based on your study habits for this course.	
15. When I study for this class, I pull together information from different sources, such as lectures, readings, and discussions.	1 2 3 4 5 6 7
16. I try to relate ideas in this subject to those in other courses whenever possible.	1 2 3 4 5 6 7
17. When reading for this class, I try to relate the material to what I already know.	1 2 3 4 5 6 7
18. When I study for this course, I write brief summaries of the main ideas from the readings and the concepts from the lectures.	1 2 3 4 5 6 7
19. I try to understand the material in this class by making connections between the readings and the concepts from the lectures.	1 2 3 4 5 6 7
20. I try to apply ideas from course readings in other class activities such as lecture and discussion.	1 2 3 4 5 6 7
21. When studying for this course, I often try to explain the material to a classmate or a friend.	1 2 3 4 5 6 7
22. I try to work with other students from this class to complete the course assignment.	1 2 3 4 5 6 7
23. When studying for this course, I often set aside time to discuss the course material with a group of students from the class.	1 2 3 4 5 6 7

Appendix E

MSLQ Pretest Questionnaire
Chinese Version

問卷說明: 以下是一份有關學習動機與策略的問卷。答案因人而異，完全沒有對與錯的區別，只要盡量依照您個人狀況回答即可。如果您覺得該題的敘述與您的狀況完全符合，圈選 7; 如果您覺得該題的敘述與您的狀況完全不符合，圈選 1。其它狀況請在 1 和 7 之間選擇描寫您最貼切的號碼。本問卷僅供研究之用，授課老師不會看到這些答案，更不會影響到您的學業成績。	完全不符合	很不符合	偶爾符合	有時符合	經常符合	大致符合	完全符合
[第一部分] 請依照您**目前對本門課的看法**，回答 1-14 題:							
1. 我想我可以把這門課裡學到的東西應用到其他課程裡。	1	2	3	4	5	6	7
2. 學習本課程的教材，對我而言很重要。	1	2	3	4	5	6	7
3. 我對本課程的專業領域很感興趣。	1	2	3	4	5	6	7
4. 我認爲學習本課程教材對我有幫助。	1	2	3	4	5	6	7
5. 我喜歡本課程的內容。	1	2	3	4	5	6	7
6. 瞭解本課程的內容，對我而言很重要。	1	2	3	4	5	6	7
7. 我相信我會在這門課拿到優秀成績。	1	2	3	4	5	6	7
8. 我確信即使是本課程中最困難的教材，我也可以理解。	1	2	3	4	5	6	7
9. 我有信心可以明瞭本課程所傳授的基本概念。	1	2	3	4	5	6	7
10. 我有信心可以明瞭本課程老師所傳授最複雜的教材。	1	2	3	4	5	6	7
11. 我有信心可以在本課程的作業及考試中表現優異。	1	2	3	4	5	6	7
12. 我預期在本課程中表現良好。	1	2	3	4	5	6	7
13. 我確信我可以掌握本課程所傳授之技能。	1	2	3	4	5	6	7
14. 如果把本課程的困難度、老師，和我個人的技能都考慮在內，我想我在本課程中會表現良好。	1	2	3	4	5	6	7
[第二部分] 請依照您**一般的讀書習慣**，回答 15-23 題:							
15. 我在讀書時，會整合不同管道(例如老師上課講授的內容、書報，以及口頭討論)得到的資訊。	1	2	3	4	5	6	7
16. 我在學習一門課的觀念時，會盡可能把它與其它課程中學到的觀念連結起來。	1	2	3	4	5	6	7
17. 我在爲某門課進行閱讀時，會試著把閱讀的內容與我既有的知識連結起來。	1	2	3	4	5	6	7
18. 我在讀書時，會把書本的主要內容與上課中得到的觀念，寫成簡短的摘要。	1	2	3	4	5	6	7
19. 我會藉由結合閱讀與上課傳授的概念，來幫助自己了解課程內容 。	1	2	3	4	5	6	7
20. 我會試著把我從課本或指定閱讀中得到的概念，應用到該門課中的其他活動上，例如上課或討論時。	1	2	3	4	5	6	7
21. 我在準備功課時，常試著把教材解釋給同學或朋友聽。	1	2	3	4	5	6	7
22. 平常在作作業時，我試著與同班同學一起合作。	1	2	3	4	5	6	7
23. 我在準備功課時，常挪出時間與同班同學討論課程教材。	1	2	3	4	5	6	7

Appendix F

MSLQ Posttest Questionnaire
Chinese Version

問卷說明: 以下是一份有關學習動機與策略的問卷。答案因人而異，完全沒有對與錯的區別，只要盡量依照您個人狀況回答即可。如果您覺得該題的敘述與您的狀況完全符合，圈選 7; 如果您覺得該題的敘述與您的狀況完全不符合，圈選 1。其它狀況請在 1 和 7 之間選擇描寫您最貼切的號碼。本問卷僅供研究之用，授課老師不會看到這些答案，更不會影響到您的學業成績。	完全不符合 很不符合 偶爾符合 有時符合 經常符合 大致符合 完全符合
[第一部分] 請依照您<u>目前對本門課的看法</u>，回答 1-14 題:	
1. 我想我可以把這門課裡學到的東西應用到其他課程裡。	1 2 3 4 5 6 7
2. 學習本課程的教材，對我而言很重要。	1 2 3 4 5 6 7
3. 我對本課程的專業領域很感興趣。	1 2 3 4 5 6 7
4. 我認爲學習本課程教材對我有幫助。	1 2 3 4 5 6 7
5. 我喜歡本課程的內容。	1 2 3 4 5 6 7
6. 瞭解本課程的內容，對我而言很重要。	1 2 3 4 5 6 7
7. 我相信我會在這門課拿到優秀成績。	1 2 3 4 5 6 7
8. 我確信即使是本課程中最困難的教材，我也可以理解。	1 2 3 4 5 6 7
9. 我有信心可以明瞭本課程所傳授的基本概念。	1 2 3 4 5 6 7
10. 我有信心可以明瞭本課程老師所傳授最複雜的教材。	1 2 3 4 5 6 7
11. 我有信心可以在本課程的作業及考試中表現優異。	1 2 3 4 5 6 7
12. 我預期在本課程中表現良好。	1 2 3 4 5 6 7
13. 我確信我可以掌握本課程所傳授之技能。	1 2 3 4 5 6 7
14. 如果把本課程的困難度、老師，和我個人的技能都考慮在內，我想我在本課程中會表現良好。	1 2 3 4 5 6 7
[第二部分] 請依照您<u>在這門課的讀書習慣</u>，回答 15-23 題:	
15. 我在讀這門課時，會整合不同管道(例如老師上課講授的內容、書報，以及口頭討論)得到的資訊。	1 2 3 4 5 6 7
16. 我盡可能把這門課的觀念與其它課程中學到的觀念連結起來。	1 2 3 4 5 6 7
17. 我在爲這門課進行閱讀時，會試著把閱讀的內容與我既有的知識連結起來。	1 2 3 4 5 6 7
18. 我在讀這門課時，會把書本的主要內容與上課中得到的觀念，寫成簡短的摘要。	1 2 3 4 5 6 7
19. 我會藉由結合閱讀與上課傳授的概念，來幫助自己了解本課程內容 。	1 2 3 4 5 6 7
20. 我會試著把我從課本或指定閱讀中得到的概念，應用到本門課中的其他活動上，例如上課或討論時。	1 2 3 4 5 6 7
21. 我在讀這門課時，常試著把教材解釋給同學或朋友聽。	1 2 3 4 5 6 7
22. 在作這門課的作業時，我試著與同班同學一起合作。	1 2 3 4 5 6 7
23. 我在讀這門課時，常挪出時間與同班同學討論課程教材。	1 2 3 4 5 6 7
	1 2 3 4 5 6 7

Appendix G

Permission for Use and Modification of the MSLQ Questionnaire

The University of Michigan

DEPARTMENT OF PSYCHOLOGY

525 E. UNIVERSITY
ANN ARBOR, MICHIGAN 48109-1109

March 3, 2005

Hui-Chuan Liao

New Orleans, LA

Dear Ms. Liao:

You are certainly welcome to use the MSLQ and to reword the questions to fit the needs of your study. We will be interested in your findings.

If you have questions or if I can help in any way, please feel free to e-mail, phone, or write. Dr. Lin, my colleague, (absygl@umich.edu) is a graduate of National Taiwan University, and I am confident that he, too, would be glad to help.

Sincerely yours,

W. J. McKeachie

Professor emeritus

Appendix H

Item Analysis on the Pilot Test for the Grammar Posttest

Item	Proportion of Students Choosing Each Response				Item Facility Index	Response-Total Correlations				r_{pbi} Discrimination Index
	A	B	C	D		A	B	C	D	
1	.24	.41*	.27	.09	.41	-.30	.17*	.12	-.03	.17
2	.12	.29	.44*	.15	.44	-.21	-.26	.38*	-.01	.38
3	.18	.09*	.21	.53	.09	-.45	.48*	.08	-.01	.48
4	.06	.12	.38*	.44	.38	-.18	-.39	.10*	.24	.10
5	.12	.03	.50	.35*	.35	-.06	-.25	.20	-.08*	-.08
6	.50*	.21	.06	.24	.50	.07*	.13	-.11	-.15	.07
7	.85*	.06	.03	.06	.85	.39*	-.08	-.01	-.51	.39
8	.50*	.06	.21	.24	.50	.34*	-.36	-.07	-.14	.34
9	.06	.71*	.00	.24	.71	-.31	-.05*	--	.22	-.05
10	.09	.24	.12	.56*	.56	-.15	-.48	.07	.44*	.44
11	.24	.15	.41*	.21	.41	-.07	-.31	.26*	-.04	.26
12	.24*	.35	.27	.15	.24	.01*	.30	-.39	.06	.01
13	.29	.09	.32	.29*	.29	.07	-.01	-.23	.16*	.16
14	.06	.09	.77*	.09	.77	.12	-.32	.25*	-.15	.25
15	.73*	.27	.00	.00	.73	.21*	-.21	--	--	.21
16	.82	.00	.06	.12*	.12	.04	--	-.03	-.02*	-.02
17	.32*	.32	.24	.12	.32	.38*	-.34	-.01	-.04	.38
18	.15	.68*	.06	.12	.68	.03	.14*	-.21	-.08	.14
19	.15	.27*	.32	.27	.27	.06	.22*	-.28	.02	.22
20	.24	.59*	.00	.15	.59	-.04	.14*	--	-.26	.14
21	.35*	.03	.35	.27	.35	.27*	-.18	-.21	.01	.27
22	.15	.27	.18*	.41	.18	-.04	-.13	.07*	.09	.07
23	.47	.18	.18*	.18	.18	-.08	.21	.21*	-.32	.21
24	.21*	.00	.32	.47	.21	.13*	--	-.35	.23	.13
25	.27	.32*	.27	.15	.32	.05	.17*	-.31	.10	.17
26	.15*	.12	.62	.12	.15	-.19*	-.15	.37	-.19	-.19
27	.35*	.35	.27	.03	.35	.32*	.05	-.39	-.04	.32
28	.44*	.15	.41	.00	.44	.17*	-.02	-.16	--	.17
29	.14*	.32	.29	.24	.14	.15*	-.33	.14	.09	.15
30	.15	.15	.28	.32*	.32	-.03	-.26	-.21	.39*	.39

(Table Continued)

Item	Proportion of Students Choosing Each Response				Item Facility Index	Response-Total Correlations				r_{pbi} Discrimination Index
	A	B	C	D		A	B	C	D	
31	.12	.50*	.15	.24	.50	-.10	.34*	-.14	-.20	.34
32	.27	.03	.50*	.21	.50	-.10	-.01	.30*	-.26	.30
33	.09	.32*	.56	.03	.32	-.11	.48*	-.36	.07	.48
34	.09	.35*	.21	.35	.35	.19	-.22*	.25	-.10	-.22
35	.47*	.06	.15	.27	.47	.31*	-.08	.11	-.28	.31
36	.06	.24	.29	.41*	.41	-.26	-.08	-.35	.51*	.51
37	.03*	.09	.74	.15	.03	.03*	-.04	.39	-.46	.03
38	.27	.06	.56*	.12	.56	-.20	-.18	.25*	.01	.25
39	.21*	.12	.18	.50	.21	.56*	.03	-.08	-.41	.56

Item	Mean	SD	r_{pbi} Discrimination Index
40	.03	.17	.21
41	.12	.33	.09
42	.12	.33	.52
43	.06	.24	.38
44	.09	.29	.52
45	.09	.39	.20
46	.18	.29	.08
47	.41	.50	.48
48	.12	.33	.07
49	.06	.24	-.03
50	.41	.50	.37
51	.12	.33	.33
52	.06	.24	.35

*Note. * denotes correct answer.*

Appendix I

Comparison Chart for the Semester Grade Rubrics of the Experimental and the Control Groups

Experimental Group		Control Group	
Mid-term exam	30%	Mid-term exam	30%
Final exam	30%	Final exam	30%
Attendance and cooperative process as evaluated by instructor	10%	Attendance	10%
Individual improvement point (IIP)	10%	Quizzes/exercises	15%
Group average improvement point (GAIP)	10%	Class participation	15%
Peer evaluation	10%		

Appendix J

Consent Form

CONSENT FORM

Title of Research Study

Effects of Different Teaching Strategies on College Students in Taiwan.

Project Director

Hui-Chuan Liao, doctoral student at the Department of Curriculum and Instruction, University of New Orleans. hcliao@uno.edu. 04-711-3709 or +1-504-280-3741.

Purpose of the Research

The purpose of this study is to examine the effectiveness of different teaching strategies in the context of an English as a Foreign Language (EFL) course being taken by college students in Taiwan.

Procedures for this Research

The study will take place in the context of the subjects' regular college EFL class. It is estimated that the procedure will take 16 weeks somewhere between March 1 to July 15, 2005, depending on the subjects' academic calendar. During the course of the study, the subjects will attend their English class as they normally would while various teaching strategies will be implemented endeavoring to enhance their English proficiency. The testing scores will be collected from the course instructor for data analysis. The data will be kept confidential, and the report of data analysis will be anonymous. At the end of the 16 weeks, the subjects will take 20-30 minutes to respond to a motivational learning strategies questionnaire. Other than what has been described above and what would normally happen or be required in a class, there would be no additional requirement for the subjects.

Potential Risks of Discomforts

There are no potential risks of discomfort other than those normally found in an English classroom. Subjects are encouraged to take part in the class activities as they normally do. If you wish to discuss any discomfort you may experience, you may email or call the Project Director listed in this form.

Potential Benefits to You or Others

During the 16 weeks of study, subjects will have opportunities to experience different teaching strategies that are intended to enhance their English ability.

Alternative Procedures

Your participation is entirely voluntary and you may withdraw consent and terminate participation at any time without consequence.

Protection of Confidentiality

Information from the examinations and questionnaires will be immediately coded by the principal investigator to protect anonymity. Only your instructor will have the pre- and posttest examination scores associated with a name. The instructor, however, will not have access to your questionnaire answers.

I have been fully informed of the above-described procedure with its possible benefits and risks and I have given permission of participation in this study.

Signature of Subject	**Name of Subject (Print)**	**Date**
	Hui-Chuan Liao	
Signature of Person Obtaining Consent	**Name of Person Obtaining Consent (Print)**	**Date**

Appendix K

Consent Form
Chinese Version

同意書

研究主題

不同教學策略於台灣大學生之效應。

研究主持人

美國紐奧良大學課程與教學系博士候選人廖惠娟。電子信箱: hcliao@uno.edu，聯絡電話:04-711-1111~3709 或 +1-504-280-3741。

研究目的

本研究旨在檢視各教學策略應用於台灣大學英文課程之效應。

研究步驟

本研究預期在 2005 年 3 月 1 日與 7 月 15 日之間，於學生正規大學英文課程中進行約 16 週，詳細時間將依學生該學期之行事曆做適度調整。在研究期間，學生一如往常參與英文課程，同時授課老師將應用不同之教學策略，幫助學生提升英文程度。學生除一般的課程規定外，同意研究人員以匿名方式採集考試成績進行資料分析，並於研究結束前填寫學習動機與策略之問卷，此外並無其他義務。

潛在之不適因素

本研究於學生正常上課之同時採集分析資料，因此，除了一般英文課程所會發生的狀況之外，本研究並無其他潛藏之不適因素。參與本研究之學生請依照平常心上課，倘若您欲討論任何於正常上課外本研究可能帶給您之不適因素，歡迎透過電子郵件或電話方式與上列研究主持人連絡。

潛在之正面因素

學生於研究期間有機會接觸各項旨在幫助學生提升英文能力之教學策略。

替代步驟

您的參與完全出於自願，您可於未來任何時候終結本同意，停止參與本研究案，絕無不良後果。

匿名之保護

研究人員在收到考卷與問卷回答時，將在第一時間以代碼刪除任何能辨識學生身分之資料，以保匿名原則。授課老師會有一份包含學生身分之考試成績，但是他不會看到您的問卷答案。

我已詳細閱讀與了解上述各項步驟以及可能之正負面效應，我在此授權同意參與本研究。

學生簽名	**學生姓名 (正楷填寫)**	**日期**
	廖惠娟	
研究主持人簽名	**研究主持人姓名 (正楷填寫)**	**日期**

Appendix L

Use of Human Subjects Approval

University Committee for the Protection of Human Subjects in Research
University of New Orleans

Campus Correspondence

Hui-Chuan Liao
Dr. John Barnitz
UNO Box 969

11/5/2004

RE: Effects of different teaching strategies on college students in Taiwan

IRB#: 08nov04

The IRB has deemed that the proposed research project is now in compliance with current University of New Orleans and Federal regulations.

Be advised that approval is only valid for one year from the approval date. Any changes to the procedures or protocols must be reviewed and approved by the IRB prior to implementation. Use the IRB# listed on the first page of this letter in all future correspondence regarding this proposal.

If an adverse, unforeseen event occurs (e.g., physical, social, or emotional harm), you are required to inform the IRB as soon as possible after the event.

Best of luck with your project!
Sincerely,

Laura Scaramella

Laura Scaramella, Ph.D.
Chair, University Committee for the Protection of Human Subjects in Research

University Committee for the Protection of Human Subjects in Research
University of New Orleans

Form Number: 08NOV04
(please refer to this number in all future correspondence concerning this protocol)

Principal Investigator Hui-Chuan Liao **Title:** Doctoral Student

Faculty Supervisor: Dr. John Barnitz ***(if PI is a student)***

Department: Curriculum and Instruction **College:** Education

Project Title: Effects of Different Teaching Strategies on College Students in Taiwan

Date Reviewed:

Dates of Proposed Project Period From 03/01/05 to 07/15/05

approval is for one year from approval date only and may be renewed yearly.

Note: Consent forms and related materials are to be kept by the PI for a period of three years following the completion of the study.

Approval Status	Date
☐ Full Committee Approval	
☑ Expedited Approval	11-5-04
☐ Continuation	
☐ Rejected	

☐ **The protocol will be approved following receipt of satisfactory response(s) to the following question(s) within 15 days:**

Committee Signatures:

Laura Scaramella, Ph.D. (Chair)
Pamela Jenkins, Ph.D.
Anthony Kontos, Ph.D.
Richard B. Speaker, Ph.D.
Gary Talarchek, Ph.D.
Kari Walsh
L. Allen Witt, Ph.D.

UNIVERSITY OF NEW ORLEANS

ALL-UNIVERSITY COMMITTEE ON THE USE OF HUMAN SUBJECTS

In order to comply with NIH policy and all federal, state and local rules and regulations concerning the treatment and use of human subjects in research, the following information must be provided. Instructional programs, internally funded and unfunded research projects using human subjects must also be approved by the Committee. Notification of the Committee's recommendations will be communicated to the investigator as quickly as possible. Please remember that research may not be initiated without approval. In the case of sponsored projects, it is therefore advisable to obtain approval at the time the project is submitted to the agency, so that the project can start as soon as the award is received. Remember that the Division of Sponsored Research will not authorize the opening of accounts unless approval has been granted by the Committee.

Form Number_____ Date Received_____

Principal Investigator: Hui-Chuan Liao **Title:** Doctoral Student

Campus Mailing Address: UNO Box 969

Department: Curriculum and Instruction

College: College of Education

Campus Telephone: 280-3741 **Home Telephone:** 280-3741

Name of Faculty Supervisor: Dr. John Barnitz
(if PI is a student)

Project Title: Effects of Different Teaching Strategies on College Students in Taiwan

Dates of proposed project period: 03/01/2004-07/15/2004

Name and address of outside funding agency to which this project has been submitted:
(if not submitted for outside funding put N/A)

N/A

Hui-Chuan Liao 10/26/2004
Signature of Principal Investigator Date

Renee Cao 10/27/04
Signature of Department Chair Date

10/27/04
Signature of Faculty Supervisor Date
(if PI is a student)

VITA

Hui-Chuan Liao was born in Taichung, Taiwan, and graduated from the University of New Orleans, earning a Master of Education degree in Curriculum and Instruction with a specialty in English Language and Literacy in July of 1999. After serving as an ESL curriculum coordinator and elementary program supervisor at Natural Way Children's School and an English instructor at Chien Kuo Institute of Technology (now Chienkuo Technology University) in Taiwan, she returned to the University of New Orleans and earned a Ph.D. in Curriculum and Instruction with a specialization in Teaching English to Speakers of Other Languages (TESOL) in December of 2005.

Liao enjoys literature and has translated eight children's books and adolescent novels. She likes to work with young people and has served as the advisor of the English Conversation Association, a college student club that aims to help members speak English fluently through communicative and contextualized activities. Liao's research interests in TESOL include cooperative and collaborative learning, learning motivation and strategies, English grammar and writing instruction, TEFL methodology, children/adolescent literature and language acquisition, teaching and learning styles, reading comprehension, listening and speaking, as well as critical discourse analysis. She is married to Tien Szu Pan, Ph.D., and together they live happily in Changhua City, Taiwan with their son, Jesse, and their Maltese daughter, Butter.