

AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON
CHINESE UNIVERSITY STUDENTS' ENGLISH
WRITING PERFORMANCE

by

XIN ZHANG

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN READING EDUCATION

2018

Oakland University
Rochester, Michigan

Doctoral Advisory Committee

John E. McEneaney, Ph.D., Chair

S. Rebecca Leigh, Ph.D.

Ledong Li, Ph.D.

Lori A. Ostergaard, Ph.D.

© Copyright by Xin Zhang, 2018
All rights reserved

To my husband, Xuelin: your steadfast love, unwavering perseverance, and constant encouragement have been the source of my strength.

To my son, Junyi, for your loving, supportive, and happy company each day.

ACKNOWLEDGMENTS

It is a challenging yet magnificent journey pursuing my study in a foreign country thousands of miles away from my home. Along the way, my sincerest gratitude goes to my dissertation chair, Dr. John. E. McEneaney, for his continual mentorship, steady support, and immeasurable feedback. My deep appreciation also goes to my committee members, Drs. S. Rebecca Leigh, Ledong, Li, and Lorri, Ostergaard for their continuous guidance, insight, and expertise. A very special thank you to Dr. Julia Smith for her patient and generous feedback to a graduate student she had never known before. With the help from each of you, I have become a stronger researcher.

I am also greatly indebted to my family members. To my dad, thank you for planting in me a seed of curiosity and exploration. To my brother, thank you for the silent support. To my father-in law, thank you for taking care of my son when I was away. To my cousins and in-laws, thank you for your collective support and encouragement.

Finally, I thank my fellow students, particularly, Lisa Hine and Laura Beth Gabrion, for your collaboration and firm friendship. I thank Oxford University Press for allowing me to republish a figure; and I thank ELSEVIER for allowing me to reuse an Essay Scoring Rubric.

Xin Zhang

ABSTRACT

AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON CHINESE UNIVERSITY STUDENTS' ENGLISH WRITING PERFORMANCE

by

XIN ZHANG

Adviser: John E McEneaney, Ph.D.

Effective writing pedagogy in higher education has been a consistent goal of researchers and instructors in the English as a second / foreign language writing practices. Formative peer feedback, a key factor in pedagogical writing practices, has been receiving growing interest (Hu & Lam, 2010). While much remains unknown regarding how the quality of peer feedback and back-feedback affects students' writing performance, this quantitative study adopted a quasi-experimental control group design to investigate primary pedagogical effects of peer feedback on university students' writing performance in an EFL context.

A total of 198 sophomores majoring in English took part in a 15-week research study. A one-way repeated measures ANCOVA analysis was conducted to examine the comparative effect between the traditional and peer feedback groups. The result revealed that beginning with slightly different writing ability, the peer feedback group significantly

made more growth in writing achievement than the traditional feedback group. Further analysis through hierarchical multiple regression analyses showed that both the quality of students' feedback and the quality of students' back-feedback were significant predictors of students' writing performance. Additionally, the quality of students' back-feedback had a slightly larger impact than that of the quality of students' feedback.

This study provides not only further evidence of the power of formative peer feedback as an effective tool to maximize learning, but also recommends the inclusion of it in the university curriculum to encourage students to direct and monitor their own learning processes and be life-long learners.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
ABSTRACT	v
LIST OF ABBREVIATIONS	xiv
LIST OF TABLES	xv
LIST OF FIGURES	xvi
CHAPTER 1	
INTRODUCTION	1
Context of Problem	1
Problem Statement	6
Definitions of Feedback	8
Significance of the Study	11
CHAPTER 2	
LITERATURE REVIEW	13
Overview	13
Theoretical Framework	14
Writing Process Theory	14
Social-constructivism	16
Self-regulated Learning	18
Input Hypothesis	20

TABLE OF CONTENTS—Continued

Output Hypothesis	23
Review of Literature	25
Issues with Teacher Feedback	25
Effects of Peer Feedback	30
Peer Training and Quality of Feedback	39
Measuring Quality and Outcomes of Peer Feedback and Back-feedback	45
Research Questions	48
Research Question 1	49
Research Question 2	50
Summary	51
CHAPTER 3	
METHODOLOGY	53
Overview	53
Design of the Study	53
Setting	54
Participants	56
Sample Size	56
Sample Characteristics	56

TABLE OF CONTENTS—Continued

Instruments	57
Feedback Guidance Sheet	57
Back-feedback Guidance Sheet	57
Essay Scoring Rubric	58
Feedback Grading Rubric	58
Back-feedback Grading Rubric	59
Pre- and Post-tests	59
Task	60
Writing Task for the Control Group	60
Writing Task for the Experimental Group	61
Peer Feedback Task for the Experimental Group	61
Training Procedure	62
Training to the Control Group	63
Training to the Experimental Group	64
Data Collection Procedure	67
Data	67
Ratings	68
Data Analyses	69

TABLE OF CONTENTS—Continued

Variables	69
Layout of the Analyses	70
Summary	73
CHAPTER 4. RESULTS	74
Preliminary Data Analysis	75
Interrater Reliability	75
Comparing Group Difference with No Covariate	75
RQ1: Comparing Instructional Approaches After Covarying English Proficiency	77
Testing ANCOVA Assumptions	78
Comparing Instructional Approaches after Taking English Proficiency into Account	79
Comparing the Impact of Instructional Approaches on the Change in Writing Performance after Taking English Proficiency into Account	80
RQ2: Relationship between the Quality of Feedback/Back-feedback and Writing Performance	83
Testing Multiple Regression Assumptions	86
Unique Impact of Feedback/Back-feedback on Students' Writing Performance	87
Examining the Change in the Impact of Feedback	88

TABLE OF CONTENTS—Continued

Examining the Change in the Impact of Back-feedback	90
Summary	91
CHAPTER 5.	
CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS	93
Conclusions and Discussions	93
The Effects of Peer Feedback on Chinese English Sophomores' Writing Performance	93
The Relationships between Quality of Peer Feedback/Back-feedback and Writing Performance	94
Implications	99
Novice Writers Benefit More	99
Well-designed Feedback Procedure	100
Wide Input	100
Limitations	102
Time Constraints	102
Non-randomized Sample	103
Monotonous Genre and Learning Context	103
Students with Low Motivation	103
Recommendations for Pedagogical Practices and Future Research	104

TABLE OF CONTENTS—Continued

Extensive and Continuous Training	104
Well-developed Instruments	105
Replicated Studies in Diversified Research Contexts	106
In-depth Exploration into Sub-aspects	106
Small Scale Qualitative Studies	107
Longitudinal Studies	107
Relationship between Peer Review and Self-review	107
Summary	108

APPENDICES

A. IRB Approval Letter	110
B. Guidance Sheets	113
C. Scoring Rubrics	118
D. Pre-and Post-test Prompts	123
E. The Research Schedule	126
F. One Writing Cycle for Experimental Group	128
G. Samples	130
H. Figures	136
I. Copyright Permission	139

TABLE OF CONTENTS—Continued

REFERENCES	145
------------	-----

LIST OF ABBREVIATIONS

ESL	English as a second language
EFL	English as a foreign language
L2	Second language
L1	Native language
SLA	Second language acquisition
SRL	Self-regulated learning
ZPD	Zone of proximal development
RQ	Research question
TEM	Test for English Majors

LIST OF TABLES

Table 3.1	Variables for ANCOVA	71
Table 3.2	Variables for Multiple Regression	71
Table 4.1	Comparing Means Between the Traditional Feedback and Peer Feedback Groups Before Intervention	77
Table 4.2	Comparing Means Between the Traditional Feedback and Peer Feedback Groups After Taking L2 Into Account	80
Table 4.3	Comparing Means Between the Traditional Feedback and Peer Feedback Groups After Taking L2 and Initial Writing Difference Into Account	82
Table 4.4	Effect of the Quality of Students' Feedback/Back-feedback on Students' Writing Performance, Controlling for Initial Writing Ability	88
Table 4.5	Change in the Effect of the Quality of Students' Feedback on Writing Performance, Overall and then Controlling for Initial Writing Ability and the Quality of Students' Back-feedback	89
Table 4.6	Change in the Impact of the Quality of Students' Back-feedback on Writing Performance, Overall and then Controlling for Initial Writing Ability and the Quality of Students' Feedback	91

LIST OF FIGURES

Figure 2.1	Output and Second Language Learning.	23
Figure 4.1	Comparing the Change Over Time in Students' Writing from Pre- to Post-treatment between Treatment and Control Groups.	83
Figure H-1	Normal P-P Plot of Regression Standardized Residual.	137
Figure H-2	Scatterplot.	138

CHAPTER ONE

INTRODUCTION

Tell me and I forget. Teach me and I may remember. Involve me and I learn.

Commonly attributed to Benjamin Franklin

Context of Problem

Writing is central to improving education (Graham & Harris, 2013). Currently, with an increasing number of universities and companies using writing to assess applicants' qualifications (Graham & Perin, 2007), writing skills, at least partially, determine an individual's academic and professional success in all disciplines (Cho & Schunn, 2007).

However, writing is one of the most challenging skills in language learning. Nunan (2001) asserts that, "In terms of skills, producing a coherent, fluent, extended piece of writing is probably the most difficult thing there is to do in language. It is something most native speakers never master" (p. 271). For university students who learn English as a foreign/second language (EFL/ESL), writing is an even more arduous task. Fully aware of the difficulty these students face, EFL/ESL writing instructors have been making continuous efforts to enhance students' writing ability and develop them into autonomous writers beyond the university.

In the process of exploring effective approaches of teaching and learning, many second language (L2) writing instructors have acknowledged the value of formative

feedback and viewed it as “crucial for both encouraging and consolidating learning” (Hyland & Hyland, 2006, p. 1). More importantly, “in higher education, formative assessment and feedback should be used to empower students as self-regulated learners” (Nicol & Macfarlane-Dick, 2006, p. 199). In other words, formative feedback aims to monitor students’ learning progress and reinforce their self-learning ability by actively involving them in the evaluation process. In higher education, though, formative feedback has still been seen as the exclusive role and responsibility of teachers (Nicol & Macfarlane-Dick, 2006), which inevitably results in multiple problems for both teachers and students.

Some of the most widespread problems are institutional issues that make formative teacher feedback impossible. The current rapid expansion of universities and colleges leads to large class sizes and teacher shortages. This tendency is more serious in countries with large populations, such as China. Consequently, teachers, and writing teachers in particular, are facing increasingly heavy workloads, which greatly influence their feedback practices. For instance, Ferris, Brown, Liu, and Stine (2011) warned that “time was a significant issue that constrained all of their [writing teachers’] response practices” (p. 222). Under such conditions, many teachers are only able to grade students’ writing in a holistic way. “Yet, holistic grading provides students with little feedback, and even holistic grading can be excessively time-consuming in large classes” (Kellogg, Whiteford, & Quinlan, 2010, p. 174). Due to the issues resulting from “larger institutional problems”

(Ferris et al., 2011, p. 223), students usually have little opportunity to receive sufficient and effective formative teacher feedback to improve their writing.

In addition to insufficient amounts of feedback, the content of teacher feedback tends to limit its effectiveness. The literature documents that most writing teachers' written feedback narrowly focuses on mechanics, such as spelling, punctuation, and grammar, rather than content aspects (Ferris et al., 2011; Lee, 2014; Liu & Hansen, 2002; Montgomery & Baker, 2007). The painful fact is that, "such form-focused feedback and revision would be unlikely to have any positive effect on the students' writing development" (Hyland, 2003, p. 228). Even when students receive feedback, the content of the feedback may not facilitate the enhancement of students' writing performance.

Another serious issue, and perhaps the greatest demerit of teacher feedback, is a lack of interaction between teachers and students, which turns students into "passive recipients" (Lee, 2014, p. 204). Owing to limited time and large ratios between teachers and students, teachers typically write their feedback directly on students' papers without asking what the students intend to express. This one-way feedback practice ends in students' passive acceptance of teacher feedback without truly understanding its significance (Lee, 2014; Zhao, 2010). Although teachers spend a large amount of time grading and commenting on students' writing, the students keep on committing the same mistakes again and again (Yang, Badger, & Yu, 2006). According to Driscoll (2011), "If students are unable to apply practices, skills, and knowledge gained in one context to a new context, they have

not truly learned and may continue to struggle in each new learning situation” (p. 1). In light of this, Lee (2014) calls for new interventions that can replace the conventional teacher feedback approaches in EFL contexts, so that feedback can truly mediate student learning.

Finally, in classes where teacher feedback dominates, students usually receive feedback from only one teacher, which greatly constrains the diversity of perspectives of complicated issues to which students in higher education should be exposed (Nicol & Macfarlane-Dick, 2006). If students fail to see a problem from different angles, they are less likely to develop critical thinking abilities needed for future careers beyond the university.

Urgently, in higher education, there is a need for a more effective approach through which students receive richer feedback, engage in bidirectional communications to exchange ideas and to increase their understanding of comments received, and advance their writing and learning abilities. More critically, this approach should empower students as independent learners in life outside of the university. Formative peer feedback is such a way of achieving effective university teaching and learning (Nicol & Macfarlane-Dick, 2006; Topping, 1998; Topping, Smith, Swanson, & Elliot, 2000), which has proven more fruitful in writing instruction (Topping, 1998).

Given that there are many more students than teachers, students have greater opportunities to receive more thorough and immediate feedback from student reviewers

(Rollinson, 2005). Unlike course teachers, who are supposed to provide feedback to multiple students within a limited time, student reviewers are usually responsible for the essays of one or several students. This allows them to spend more time on reviewing and producing substantive feedback.

Peer feedback activities tend to be bidirectional and interactive (Rollinson, 2005; Zhao, 2010). During interactions with each other, students seek clarifications, exchange ideas about informational gaps, negotiate on irrelevant and insufficient information, make suggestions, and reflect upon comments. These activities are not only socially, but also cognitively and meta-cognitively, demanding, which promotes the exchange of ideas among students, polishes their social skills, and ultimately consolidates their writing and learning abilities, thus developing them into independent learners.

Peers, particularly those in higher education, are a rich source of feedback. Being adult students, they are capable of offering unique, individualized, and diversified responses to their peers' essays (Rollinson, 2005). Receiving feedback from multiple student reviewers is more likely to broaden student writers' horizons by exposing them to various perspectives on the same issue. This diversity of perspectives enables them to look at phenomena or problems from different angles, which will refine their critical thinking abilities and develop them into lifelong autonomous learners.

Problem Statement

Since the introduction of formative peer feedback into EFL contexts, a growing number of prestigious universities in Hong Kong, Taiwan, and Mainland China have incorporated this approach into their writing curriculum. Unfortunately, the university for which I have been working has not yet recognized the value of formative peer feedback, a deficiency that is all too common among mid-level Chinese universities. This observation is consistent with the general statement by Nicol and Macfarlane-Dick (2006), who found that, in the typical university, formative feedback is still seen as the exclusive responsibility of instructors. This view, coupled with the heavy course load of instructors, makes traditional writing instruction the typical approach in these universities. In my university specifically, it is common for writing instructors who teach English majors to be assigned at least four classes per term (approximately 120 students in total), which makes it impossible for them to offer constructive written feedback on individual students' essays. Teachers usually select several volunteers' essays from each class and provide detailed written comments on these samples. They then use these essays to demonstrate and guide discussions on the major issues that might be shared by the whole class. During the class periods, teachers make every effort to provide oral feedback on each student's essay. Throughout the entire term, teachers may assess and produce written comments on an individual student's assignments three or four times, but time constraints preclude them from doing so for each assignment. As a result, although instructors invest a lot of

time, there is little improvement in students' writing performance. This makes writing instruction in China a frustrating practice indeed (Yang et al., 2006).

I have been teaching writing at this university for almost four years, during which time I have been exploring a possible solution to this long-existing problem in writing education. Empirical research and personal experience support the potential solution of formative peer feedback, which has been widely used in famous universities in China and has proven effective in writing instruction in past decades (Ge, 2011; Ho, 2015; Min, 2005, 2006, 2008; Zhao, 2014). The literature on peer feedback shows that a variety of studies have been conducted, covering the comparative effects of teacher feedback and peer feedback (Yang et al., 2006; Zhao, 2010), peer training and writing performance (Min, 2005, 2006, 2008; Zhao, 2014), comparative effects of traditional face-to-face and online peer feedback (Ho, 2015), effects of online feedback (Ge, 2011; Liang, 2010), and decision-making in the peer review process (Ma, 2010).

Surprisingly, little quantitative research has been conducted to examine the correlations between the quality of peer feedback and students' writing performance. In addition, although peer feedback generally includes both reviewing and revising activities, existing studies overwhelmingly focus on the relationship between feedback and writing performance, leaving the relationship between back-feedback (the response of writers to reviewers' feedback) and writing performance unexamined. Much remains to be

discovered regarding the relationship between students' perception of peer comments received and the receivers' writing performance. This study aims to fill the void in the existing literature.

Definitions of Feedback

Provided this study involves different types of feedback, definitions and analyses of these types of feedback are presented first, followed by definitions of the key terms for this study.

Feedback, also referred to as assessment, comment, evaluation, review, and response, is generally defined as “input from a reader to a writer with the effect of providing information to the writer for revision” (Keh, 1990, p. 294). In the field of academic writing specifically, the function of feedback is to raise writers' awareness of the informational, rhetorical, and linguistic expectations of target readers (Hedgcock & Lefkowitz, 1994). In other words, feedback intends to help a writer be conscious of where he or she has confused or misled readers with insufficient information, a lack of development of ideas, illogical organization, or inappropriate tense and choice of words.

In terms of form, nature, and source, feedback is broadly categorized into written or oral; summative or formative; teacher or peer feedback. The definitions are presented as follows:

Written feedback refers to handwritten or typed comments provided directly on the students' writings (Ferris, 1997).

Oral feedback refers to verbal discussion of a student's paper with the writer (Vorobel & Kim, 2014).

Summative feedback refers to written end comments on final drafts (Hyland & Hyland, 2006). This type of feedback usually includes a final quantitative mark and grade, aiming to evaluate the student's success or failure and allowing few chances for further improvement (Topping et al., 2000).

Formative feedback can be oral, written, or a combination of both (Topping, 1998). It usually involves a process of assessing, generating feedback, receiving feedback, and revising. It provides rich and constructive qualitative evaluation of writers' strengths and weaknesses. More often than not, it has been used as a method to maximize learning rather than as a tool to evaluate students' performance (Topping et al., 2000).

Teacher feedback refers to "all responses that a teacher makes on a student's draft including shorthand symbols, punctuations markers, grade earned, and in-text as well as end comments" (Lee & Schallert, 2008b, p. 507). This definition stresses the requirement for written feedback on every student's assignment but ignores the oral feedback teachers may provide to students.

Based on previous work by Tsui and Ng (2000) and Zhu (2001), peer feedback is defined by Hu (2005) as "a collaborative activity involving students reading, critiquing and providing feedback on each other's writing, both to secure immediate textual improvement and to develop, over time, stronger writing competence via mutual

scaffolding” (p. 321–322). By claiming that peer feedback is interactive and formative by nature, this emphasis promotes both short-term and long-term learning.

Despite the clarity, this definition misses one of the key components of peer review activities: revision. According to the revising model proposed by Flower, Hayes, Carey, Schriver, and Stratman (1986), reviewing includes two sub-processes: evaluation and revision. Evaluation followed by reflection upon one’s own work facilitates better learning. Although this model was initially developed to understand writers’ revisions of their own papers, it sheds light on the peer review process as well. Critically evaluating and incorporating feedback received is another essential component in peer review activities (Li, Liu, & Zhou, 2012).

Based on the definitions above, the key terms in this study, namely traditional feedback, peer feedback, and peer back-feedback, are provided as follows:

Traditional feedback is hereby defined as teachers’ oral feedback on major issues shared by the whole class, usually through commenting in class on several student volunteers’ essays. Traditional feedback also includes teachers’ written responses to students’ writing, but neither oral nor written feedback on every assignment of every student is guaranteed. The difference between teacher feedback and traditional feedback is that the former provides written feedback on every assignment completed by every student.

Peer feedback refers to a sequence of oral comments (in native language) and written comments (in English), all of which are provided to a partner by a paired student.

Peer back-feedback refers to student writers' responses to student reviewers' feedback. It occurs in the same sequence as that of peer feedback. By producing back-feedback, student writers assess their own writing through reflecting upon peer feedback received and improve their writing with follow-up revision activities. The nature of both peer feedback and back-feedback is formative and interactive, aiming to yield short- and long-term learning outcomes and skills.

Significance of the Study

This study has the potential to provide institutional decision-makers with some insights into the application of formative peer feedback as a tool to consolidate students' knowledge acquisition and to develop their learning autonomy in EFL contexts. As institutional contexts tend to determine students' learning strategies and attitudes (Gan, 2009), this study may shed prescriptive light on future policies for many academic institutions that have largely ignored the power of formative peer feedback (Liu & Carless, 2006).

This study may have significance to the theoretical perspectives of formative peer feedback. The majority of previous studies on peer feedback have laid their theoretical ground on social constructivism, centering on the joint construction of knowledge through collaboration and group work (Lundstrom & Baker, 2009; Topping, 1998; Zhao,

2010). This study draws attention to the theoretical framework of peer feedback by viewing it as a self-regulated process, during which students actively monitor their learning process through reviewing and revising activities.

Findings of this study may make contributions to the pedagogical implications for implementing peer feedback as a powerful learning tool in EFL contexts. It may not only open doors for those teachers who have not yet recognized the value of formative peer feedback, but also provide practical guidelines for teachers who intend to explore how to incorporate peer feedback into their curriculum.

CHAPTER TWO

LITERATURE REVIEW

Overview

Peer feedback – that is, students’ responses to their peers’ writing – has been widely used as a learning tool in many institutions for more than fifty years (Sluijsmans, Brand-Gruwel, & vanMerrienboer, 2002). The essential assumption underlying this study is that, with extensive training, peer feedback has the potential to be an important alternative to teacher feedback in higher education ESL/EFL writing instruction settings. Peer feedback can be used as a learning method to facilitate students’ writing performance.

The purpose of this study is to ascertain the effect of peer feedback on Chinese EFL university students’ writing performance. Specifically, it investigates the effects of peer feedback versus traditional feedback; it addresses the relationship between the quality of peer feedback and students’ writing performance, including the relationship between the quality of peer back-feedback and students’ writing performance.

This chapter reviews literature relevant to the study’s research questions and covers three sections: theoretical framework, review of literature relevant to the research questions, and proposal of research questions. The first section provides the theoretical framework of peer review which entails 1) writing process theory 2)

social-constructivism 3) self-regulated learning 4) input hypothesis, and 5) output hypothesis. The second section reviews previous studies on teacher feedback and peer feedback. It covers 1) issues with teacher feedback, 2) effects of peer feedback, 3) review of peer training and quality of feedback, and 4) measuring quality and outcomes of peer feedback and back-feedback. The purpose of this section is to provide an overview of the issues that have been studied so far. The third section addresses the research questions and the rationales for each.

Theoretical Framework

Diverse theoretical perspectives have been proposed as foundations for examining peer feedback. This section discusses the five theories listed in the aforementioned overview respectively.

Writing Process Theory

Before the introduction of Process Approaches, Product Approaches have dominated second language writing instruction for decades (Zamel, 1982). Product Approaches view writing as a linear process which is “mainly concerned with knowledge about the structure of language, and writing development as mainly the result of the imitation of input” (Badger & White, 2000, p. 154). In a typical product-orientated writing class, repetitive drill training on rules and rhetoric patterns often occurs, which leads to the perception that form is more important than meaning (Zamel, 1985).

In the late 1960s and early 1970s, Process Approaches first appeared in native language (L1) writing instruction and then brought about great changes in second language (L2) writing pedagogy (Zamel, 1982). Unlike Product Approaches, Process Approaches regard writing as a recursive process, during which meaning is constructed and ideas are generated, organized, and refined (Zamel, 1982). In addition, Process Approaches stress the social nature of writing, viewing it as “internalized social talk made public and social again” (Bruffee, 1984, p. 641). When writing is seen as a production of the interaction of writer, audience, texts, context, and language, it lays ground for collaborative learning activities in which peers’ potentials as audience and consultants are discovered (Farris, 1987).

More importantly, the shift of focus from product to process leads to the role of feedback changing from a summative to a formative position that allows for growth in skills possible (Hyland & Hyland, 2006). In a product-oriented class, students typically write only one draft for which the teacher offers general comments on linguistic issues or simply provides a numerical score, leaving students with minimum information on the quality of the writing and little or no opportunity for revision. By comparison, students are encouraged to write multiple drafts in a process-oriented class. Among drafts, they receive comments from the teacher and peers on the various stages of brainstorming, drafting, revising, and editing- and where possible, publishing. Students have

opportunities to reflect upon and revise their drafts, making their writing performance more subject to improvement.

Although initiated in L1 instruction, researchers generally agree that the composing process is similar in L1s and L2s: it is a process of discovering meanings and ideas (Zamel, 1982). Multiple drafts of prewriting, writing, and revision have been gaining more and more importance in L2 writing classes, which lays a solid foundation for research on the value of peer feedback.

Social-Constructivism

Social constructivism also strongly informs peer review activities. According to Vygotsky (1978), the representative social constructivist, learning is a cognitive activity occurring in social interaction rather than an individual and isolated activity. At the center of learning is the concept of the Zone of Proximal Development (ZPD): the distance between what a child can do independently and what he/she can achieve with external help. Vygotsky claimed that within children's ZPD, there are functions "not yet matured" but "in the process of maturation," functions that could be termed the "buds" or "flowers" rather than the "fruits of development" (Vygotsky, 1978, p. 86). These functions or skills occur first in interactions with people who are more capable and then within the individual child. Even though such interactions are written and asynchronous, these new skills are taken over by the individual child and internalized (Davydov & Kerr, 1995).

Vygotsky's ZPD was initially proposed to promote children's learning and development which has obvious limitations in modern higher education. Wells (1999) expanded the concept, which powerfully supports adults' learning.

Wells' first contribution to expanding one's understanding of Vygotsky's ZPD is that human beings are life-long learners whose cognitive processes continue to develop throughout one's life. People of all ages and at various stages of their lives experience ZPD, not simply children.

His second contribution is the interpretation of external assistance. Rather than assistance from a more capable individual, external assistance is often obtained from peers and many other sources. Peers are good sources of external assistance because most collaborative activities involve various tasks and it is less likely for a group member to excel beyond the others in all respects. Furthermore, it often happens that the group as whole, in which no member is more capable than his or her peers, is able to construct a solution that none could have achieved alone. Additionally, in the information age, assistance does not necessarily have to come from human beings solely. The internet and books provide great resources as well.

Wells' third contribution is the dynamic ZPD. Unlike Vygotsky (1978) who stated that there is an upper bound for the individual's ZPD at a certain point of time, Wells argued that ZPD is not the fixed attribute of the learners; rather, "it is treated as an attribute, not of the student alone, but of the student in relation to the specifics of a

particular activity setting” (Wells, 1999, p. 318). Put another way, the ZPD is determined by the interaction between learners as they take part in a particular activity. In principle, the upper bound depends on both the way the interaction unfolds and on any independent estimate of the current potential of the participants. Hence, ZPD is a relative dynamic zone and will change with the development of the learners’ ability.

Peer feedback in writing activities strongly matches social-constructivism by providing students multiple opportunities to co-build knowledge and skills within each other’s ZPD. Interaction with different individuals enables them to explore their potential in different areas. Whether their peers are more intellectually capable in producing feedback or not, they can achieve their individual goals and extend their mastery of certain knowledge such as linguistic awareness, writing skills, and critical thinking.

Self-regulated Learning

Self-regulated learning (SRL) is one of the diverse theoretical foundations of peer review activities. Zimmerman (2008) referred to self-regulated learning as self-generated thoughts, feelings, and behaviors that are systematically oriented toward the attainment of individual learning goals. SRL “emphasizes autonomy and control by the individual who monitors, directs, and regulates actions toward goals of information acquisition, expanding expertise, and self-improvement” (Paris & Paris, 2001, p. 89). SRL does not mean learners solely decide their learning goals; rather, the teacher usually designs the learning task and determines the evaluation standards (Nicol & Macfarlane-Dick, 2006).

SRL features deep engagement in cognitive and meta-cognitive tasks and activities. Self-regulated learners, with the help of teachers, set goals and then actively monitor their learning progress by observing their current performance, comparing their performance to the criteria and the goals that have been established, reacting and responding to perceived differences between their current level of performance and desired standards, and then take further actions to improve their performance (Nicol & Macfarlane-Dick, 2006).

Rather than learning in isolation, self-regulated students place importance on interactive work with and effective assistance from others, such as teachers, peers, and friends. During their learning process, self-regulated learners assess their progress by critically interpreting both internal and external feedback in relation to their goals (Nicol & Macfarlane-Dick, 2006). Feedback helps them confirm existing information, add new information, replace incorrect information, and tune conditional application of information (Butler & Winne, 1995). This process enables students to evaluate their current state of learning and the goals set at the beginning and to react in a way that closes the gap to achieve them.

Similar to self-regulated learning, peer review which is regarded as a formative developmental process (Mulder, Pearce, & Baik, 2014), involves multiple cognitive, meta-cognitive, and interactive tasks and activities. At the initial stage of the process, in order to provide constructive feedback, students develop a clear understanding of the task

and the criteria the teacher selects. Next, they review their peers' writing, identify their strengths and areas of need, generate feedback, provide solutions, and facilitate their peers' achievement of their learning goals. Finally, students receive feedback on their own product. They often critically reflect on the feedback, compare their peers' products with their own, evaluate strengths and weaknesses of their own products, and then take actions to improve their own products.

All these activities center on feedback which is “an inherent catalyst” (Butler & Winne, 1995, p. 246) for self-regulated skills. While working on these activities, students are aware of their specific goals. Next, they execute strategies to meet the goals while applying meta-cognitive monitoring to determine whether the goals or strategies require adaptation within the process. Apparently, well-designed formative peer feedback can be used to enhance students' self-regulated skills and “to empower students as self-regulated learners” (Nicol & Macfarlane-Dick, 2006, p. 199) in their life time.

Input Hypothesis

Theoretical support for the application of peer review in ESL/EFL writing instruction is also found in second language acquisition (SLA) theories, including the Input Hypothesis (Krashen, 1981) and the Output Hypothesis (Swain, 1985).

Many researchers such as Long (1983b) and Hutten locher, Haight, Bryk, Seltzer and Lyons (1991) have argued that input plays a critical and necessary role in language

acquisition. Krashen (1981) put forth the idea of the Input Hypothesis, strongly arguing that the availability of comprehensible input is a prerequisite for SLA.

The Input Hypothesis proposes two concepts—language learning and language acquisition, which are two separate processes that internalize the target language. In the language learning process, students consciously focus on forms and rules, while being fully aware that they are learning a language. Conversely, in acquisition, learners are conscious that they are using the language to communicate without the awareness that they are learning the language. Acquisition is a subconscious and intuitive process, similar to the way children pick up their native language. Krashen’s research suggests that acquisition is more responsible for language fluency than learning is. Comparatively, greater engagement in acquiring activities is more likely to yield better language performance.

In Krashen’s Input Hypothesis, two conditions need to be satisfied for acquisition to occur. One is comprehensible input ($i+1$) and the other is low anxiety. In comprehensible input, “ i ” refers to learners’ current language competence, and “ 1 ” refers to their next higher level of competence. Comprehensible input is essential but not sufficient for the acquisition process to take place.

The second condition for acquisition is low anxiety, which makes comprehensible input possible. Students with high anxiety typically absorb less input; even though they

understand the message, the possibility for them to internalize the information is strikingly reduced.

In process-oriented writing classes, peer review tasks provide students with more opportunities to be involved in acquiring activities, such as clarification, explanation, and meaning negotiation rather than in work focused on linguistic forms, rules, and error corrections. Additionally, peer interaction is more likely to produce comprehensible input ($i+1$), because students often share similar educational backgrounds and life experiences. It is highly possible that the feedback they receive offers meaningful expressions beyond their current competence level.

Peer review activities also satisfy the second condition of low stress by creating a student-friendly environment which facilitates acquisition. In traditional writing classes, teachers' overemphasis on error-free texts usually frustrates students and raises their anxiety, which inevitably impedes their acquisition process. In opposition to the traditional class, peer collaboration encourages students to discover their intended meanings and to express themselves more confidently through their writing. More importantly, if students feel less anxious and embarrassed about errors, they are able to concentrate more on meaning-negotiation instead of language itself. Hence, the Input Hypothesis is theoretically supportive of the application of peer review in EFL contexts.

Output Hypothesis

Given that not all the input that students are exposed to has the potential to become output, many researchers argue that comprehensible input alone is not enough for SLA. Swain (1985) proposed the Output Hypothesis and refined it in her later works (Swain &

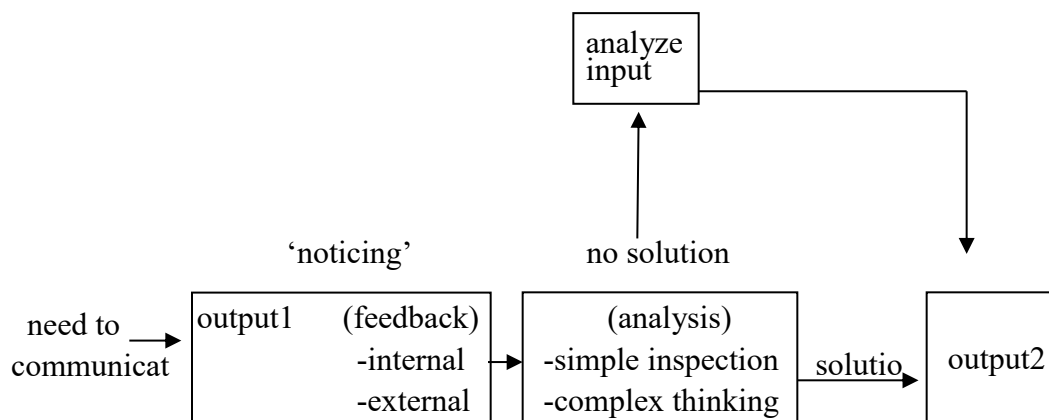


Figure 2.1 Output and second language learning (Swain & Lapkin, 1995, p. 388)

Lapkin, 1995). The Output Hypothesis emphasizes the noticing function in the process of producing L2. The noticing function holds that, while producing the target language in writing, students may be aware that they do not know how to precisely express their intended meaning, which “forces” students to modify their output. What occurs between the output and its modified form is part of the process of SLA (see Figure 2.1).

Similar to Swain, Schmidt (1990) proposed the Noticing Hypothesis, stating that (1) not all input will receive equal attention and (2) only the input noticed will become available for further effective processing. In other words, students give selective attention to certain input, and only the noticed part will be meditated upon.

As Figure 2.1 shows, in process-oriented writing classes, noticing is the attention that a teacher or student reviewers deliberately draw to certain issues in their peers' papers, which redirects students to focus on these problems and find solutions to them. The composition process occurs in the following order. First, the students need to communicate in written forms, which leads to the first output—or the first drafts. After the first drafts, student writers receive some form of feedback – either internal (self-feedback) or external (teacher/peer feedback) – which conveys to the writers that certain ideas they produced are problematic. When problems are diagnosed, students further inspect them and find solutions in their second drafts. If they cannot find solutions, they may resort to input with more focused noticing until they reach a solution. Throughout this process, the writers' previous output together with their feedback serves as the input for subsequent drafts, during which noticing acts as a mediator between outputs and inputs, triggering mental processes that lead to effective learning.

Qi and Lapkin (2001) further argued that although the composition process promotes significant noticing, the quality of that noticing is a more vital issue in L2 writing pedagogy. Peer assessment serves as an effective way to promote quality of noticing.

Compared to teacher feedback, peer feedback tends to be more immediate and individualized. Additionally, rich negotiation and deep discussion between peers draw the writers' attention to selected problems, thus increasing the possibility of quality noticing. Consequently, the writers become more alert to these problems in their own writing and will then attempt to find solutions in subsequent drafts. Therefore, Output Hypothesis serves as a solid foundation for peer review activities.

Review of Literature

Since this current study was conducted with Chinese EFL university students, the review of literature focused on peer feedback research in higher education in ESL/EFL contexts.

Issues with Teacher Feedback

Numerous studies concerning writing instruction suggest that students demonstrate preference for teacher feedback over peer feedback (Ren & Hu, 2012; Yang et al., 2006; Zhao, 2010). However, a large body of research has revealed problems with teacher feedback (Lee & Schallert, 2008a, 2008b; Yang et al., 2006; Zhao, 2010). These issues are examined based on literature about comparative effects of teacher and peer feedback, students' perception of and responses to teacher feedback, and teachers' self-assessment of their feedback practices.

Connor and Asenavage (1994) examined the impact of peer and teacher feedback on eight ESL students from different countries at a university in the US. Interestingly, they observed that about 60% of the revisions were the result of neither peer feedback nor teacher feedback; instead, they were triggered by self-feedback. In contrast, teacher feedback triggered only 35% of the revisions, which addressed primarily surface issues.

In line with Connor and Asenavage (1994), Paulus' (1999) comparative study with 11 undergraduate ESL students at a public university in the US yielded similar results. Only 34.3% of revisions in the participants' essays resulted from teacher feedback. On the contrary, a high percentage (51.8%) of revisions were attributed to other sources, including self-revision. The large percentage of self-revision in both studies may suggest a lack of trust in teacher feedback, which would be discussed later in Lee and Schallert's (2008a; 2008b) studies.

Differing from the above two studies, Yang et al.'s (2006) study with 12 Chinese students reported a high incorporation rate of teacher feedback (90%). The rate of peer feedback was 76%. Although participants viewed teacher feedback as authoritative, most teacher-influenced revisions happened at the surface level. Conversely, peer feedback triggered more content changes and "slightly more successful revisions" (p. 189). Furthermore, peer feedback was associated with a greater degree of student autonomy.

Unlike the comparative studies discussed above, which applied frequency measures to examine peer and teacher feedback, Zhao (2010) uniquely used stimulated recall

interviews to examine 18 Chinese students' use and understanding of peer and teacher feedback. Using the feedback on the students' first drafts and the revisions in their redrafts as the stimuli, she conducted these interviews in the participants' L1 by asking them three questions:

1. Could I help you with any feedback instance that you felt confused about?
2. I noticed you changed this language point here. What were you thinking when you changed it?
3. I noticed you deleted this part. What were you thinking when you deleted it? (p. 9)

The feedback instances that the students mentioned in their responses to the first question were labeled as not understood. In their responses to the second and third questions, the feedback instances in which the students provided justifying reasons were labeled as understood; otherwise, they were labeled as not understood. The findings uncovered that 83% of peer feedback comments adopted in participants' revisions were fully understood. In contrast, only 58% of teacher responses were incorporated with a real understanding of their value. This study provided an explanation of a long-lasting phenomenon: that even though teacher feedback was highly appreciated, "the mistakes in these essays keep on repeating themselves" (Yang et al., 2006, p. 180). This study also unlocked the potential power of peer feedback because understanding is the prerequisite of true learning.

Considering the need for individualized feedback, Ruegg (2015) conducted a longitudinal quasi-experimental research study (N=51) to investigate the effects of teacher and peer feedback on Japanese students' writing ability. Both the teacher and peers used the same feedback form to produce systematic feedback. The form included

four questions (either global or local) that were meant to be asked by the writer and answered by the reviewer and a prompt for one constructive suggestion from the reviewer. The gains between the pre- and post-tests of both groups indicated that the participants in the peer feedback group made similar writing improvements in organization, vocabulary, content, and total essay quality to those made by participants in the teacher feedback group. The findings of the study indicated that the effectiveness of teacher feedback surpassed that of peer feedback only in grammatical aspects of writing.

From a new perspective, Lee and Schallert (2008a, 2008b) examined students' responses to the teacher's written feedback in a Korean context. Lee and Schallert's (2008a) case study involved a non-native teacher and two EFL students as participants; in Lee and Schallert's (2008b) research, a non-native teacher and 14 EFL students were included. Data sources included formal, informal, and text-based interviews, class observations, and writing samples with the teacher's written comments. Both studies pointed out that the feedback and revision cycle was reciprocal, bidirectional, and social. In this interaction, trust between the teacher and students deeply influenced students' responses to the teacher's feedback and hence, positively impacted their revision and writing performance. In other words, students with higher levels of trust in their teacher responded more to these comments and made greater improvement in the quality of their written work. This may partially provide a reason why there was a high percentage of self-revision in Connor and Asenavage's (1994) and Paulus' (1999) studies.

In recent years, teacher feedback was also examined from the teachers' perspective. Montgomery and Baker (2007) surveyed both second language (L2) teachers (N=13) and students (N=98) in order to gather information about students' perceptions of teacher-written feedback and about how well teachers' self-assessments matched their actual performances. Their findings revealed that students' perceptions coordinated well with teachers' self-assessments. However, throughout the writing process, the instructors provided more feedback on local issues, such as grammar and mechanics, than was actually perceived by them.

In a similar line, Ferris et al. (2011) investigated teacher feedback purely from the teacher's perspective. They surveyed 129 ESL university writing teachers about their feedback practices and followed up with "23 teacher case study narratives" (p. 219). They found that most teachers in the sample focused "heavily on grammar or language issues in feedback - especially with L2 writers" (p. 224) and that these teachers were largely unaware of the challenges L2 writers faced. The teachers' feedback preferences were attributed to the teachers' overwhelming workload.

Summary. The effectiveness of teacher feedback is often insufficient to facilitate consistent improvement in general writing performance. These studies reveal a need for other sources of feedback, such as peer response, especially when the teacher-to-student ratio is low.

Effects of Peer Feedback

Research on peer feedback began with native speakers and was then expanded to second language learners across diversified content areas and at different ages worldwide (Crinon & Marin, 2010; Lu & Law, 2012; Mulder, Pearce, & Baik, 2014). After a comprehensive review of 109 articles (most centering on the L1 context) about peer feedback in higher education, Topping (1998) concluded that peer review of writing “appears capable of yielding outcomes at least as good as teacher assessment and sometimes better” (p. 262). Topping went on to further classify these outcomes into four categories: cognition and meta-cognition, affect, social and transferrable skills, and systemic benefits. While Topping evaluated the potential of peer feedback in a more general sense, encompassing areas such as science, reading, etc., Liu and Hansen (2002) focused solely on its effectiveness in writing improvement. They claimed that peer feedback activities in L2 writing contribute to students’ cognitive, affective, social-cultural, and linguistic growth. This section reviews studies in ESL/EFL writing and examine these outcomes.

Cognitive and meta-cognitive domain. According to Topping (2009), “cognitive and meta-cognitive benefits can accrue before, during, or after the peer assessment” (p. 23), a trend which has been shown in numerous studies. Peer feedback activities usually foster improved writing performance, higher-order cognitive abilities, writing autonomy,

and audience awareness. These benefits represent a solid foundation for students' effective writing abilities.

First, a large body of empirical studies has shown that peer review activities help to produce overall positive effects on students' writing performance. For example, Diab (2011) compared the students' writing quality of a peer feedback group and a self-feedback group at a Lebanese university, reporting that the peer feedback group produced significantly better second drafts than the self-feedback group. Kamimura (2006) revealed that Japanese university students of both high- and low-English proficiency benefited from peer review and improved their writing quality. In Ge's (2011) study, all the Chinese participants in three groups of different writing abilities obtained satisfactory results from an internet-based peer review process.

Some studies specifically pointed out the aspects where learners made more growth. Ho (2015) compared the effects of face-to-face and online peer review on 13 second year English majors' comments and revisions at a university in Taiwan of China. All the students agreed that "peer comments helped them the most with revising the global problems such as identifying irrelevant ideas, offering more supporting ideas, and organization" (p. 11). This confirms previous findings that peer feedback triggered a higher percentage of content changes (Lundstrom & Baker, 2009; Yang et al., 2006), which is considered to be signs of better writing (Liu & Sadler, 2003). Apart from improvement on global issues, reading peers' papers also contributes to students'

improvement on local aspects of writing, such as acquisition of some new words, phrases, and sentence structures (Zhao, 2014).

Even in multicultural classrooms where students' divergent notions of good writing are considered a drawback, encouraging findings were reported. Tang and Tithecott (1999) exposed that "a heterogeneous group of students in a classroom situation reap many of the same benefits of peer interaction as the students in a more controlled, homogeneous environment" (p. 34). In Zhu's (2001) study, students were divided into mixed groups of native speakers and non-native speakers. A comparison of their feedback showed that non-native students were able to produce as many global comments as the native speakers could in written feedback. A recent study by Vorobel and Kim (2014) also observed that "ESL students could not only identify the strengths and weaknesses of their peers' writing, but also point them out to the writers and usually suggest how to improve the paper" (p. 715). All the studies provided evidence that peer feedback can yield productive performance, even in context with different L1s and cultural differences.

Second, writing involves demanding higher-order cognitive abilities (Vygotsky, 1986), which can be sharpened and consolidated through peer feedback activities. The complexity of the reviewing process nurtures development of these abilities, such as critical thinking and problem-solving skills (Demiraslan Çevik, 2015; Paulus, 1999; Vorobel & Kim, 2014; Zhu & Mitchell, 2012). As reviewers, students are challenged

with tasks of identifying informational gaps, clarifying the author's intention, discussing alternative points of view, and co-constructing meaning. Feedback receivers need to explain their writing purpose, reflect upon the givers' comments, and select helpful suggestions. The active and reflective role they play in these activities ultimately expands their advanced higher-order cognitive abilities.

Third, peer review leads to development of writer autonomy. Partially because of a lack of confidence in their reviewers' ability to provide valid feedback, writers actively and critically evaluate suggestions received, which helps them to clarify their writing purpose and strengthens their ability to select good evidence (Biggs & Tang, 2007). For example, Yang et al. (2006) observed a strong tendency of learner autonomy in the peer feedback group; this finding is in line with Mendonca and Johnson (1994), Villamil and De Guerrero (1996), and Berg (1999). Rather than blindly taking their peers' feedback, students critically assessed the suggestions and selectively incorporated the valid ones into their revisions.

In addition to benefits to students' cognitive abilities, peer review activities frequently promote the students' meta-cognitive ability by cultivating their audience awareness (Zhao, 2014). Evaluating multiple papers provides student assessors with repeated opportunities to identify good or poor features in peers' texts, which sharpens their awareness of "what counts as good quality work in the subject area" (Hounsell, McCune, Hounsell, & Litjens, 2008, p. 55). Thus, in their future writing, students can

often consciously reflect upon their own papers, ask themselves how an audience would interpret their writing, and avoid informational gaps or misleading information.

Despite multiple cognitive and meta-cognitive advantages brought about by peer review, some disadvantages are also reported. A common negative effect comes from students' limited perspectives (Zamel, 1985) of good writing. Viewing peer review as “a task to correct grammar and format rather than an opportunity to exchange and discover ideas with the writers” (Min, 2008, p. 300), many students tend to focus on micro-level error correction rather than macro-level content issues in peer review (Leki, 1990). Such perspectives often result in less constructive feedback (Ge, 2011; Hu, 2005; Ma, 2010), which in turn, affect students' ability to select valid suggestions from peer feedback when revising their own work (Hu, 2005; Stanley, 1992).

Affective domain. A substantial body of studies in ESL/ EFL writing education reports that peer feedback reduces students' apprehension and increase their confidence in writing. For instance, Tang and Tithecott (1999) found that most of their ESL students (N=12), regardless of their proficiency levels, benefited from the peer review sessions and increased their self-confidence, an observation which was also confirmed by Ge (2011), whose students felt growing confidence in their own writing after the peer review session.

Peer review may also increase students' motivation through enhanced personal responsibility. In Ge's (2011) study, all 36 Chinese students were “willing to assume

responsibility” (p. 87) and thought “it was very interesting and challenging to comment on others’ writings” (p. 85). Similarly, Villamil and de Guerrero (1996) found that among the 54 participants, “most dyads established joint responsibility for accomplishing the task” (p. 68), which encouraged more rethinking and revisions (Stanley, 1992; Tuzi, 2004).

Despite the various arguments for the advantages of peer review in these domains, a variety of studies have shown negative effects of peer review. Some studies report strong preference for teacher feedback over peer feedback (Ren & Hu, 2012; Yang et al., 2006; Zhao; 2010). Other studies revealed students’ lack of confidence in their abilities as both reviewers and writers (Guardado & Shi, 2007; Hu, 2005; Zhang, 1995).

Social-cultural domain. Another area where peer feedback yields multiple benefits is the social-cultural domain. One of the salient benefits of peer feedback in this domain might be its contribution to the creation of a supportive environment for learning to write. Zhao (2014) conducted a study on teacher-supported peer review in an intermediate EFL writing class (N=18) at a university in China. Her findings showed that peer assessment has been socially supportive: 1) the knowledge gap between peers, as compared with teachers, is narrower due to shared education background and life experiences, which allows students to understand each other better regarding certain social issues; 2) the equivalence of social status between peers creates a more relaxing atmosphere than with the teacher, which fosters a more interactive and in-depth peer discussion. Many previous

empirical studies also observed that peer review is “more empathetic in perspective and more explicitly explained in a friendlier tone” (Min, 2008, p. 300), which creates “a fruitful environment” (Liu & Sadler, 2003, p. 194) and promotes engagement and interaction in learning writing (Jacobs, Curtis, Braine, & Huang, 1998; Villamil & de Guerrero, 1996).

Another significant social benefit resulting from peer feedback might be a sense of wider audience (Jacobs et al., 1998; Paulus, 1999; Zhang, 1995). Owing to the fact that there are more students than teachers in most classrooms, students are likely to have a more plentiful peer response, sometimes from varied cultures, which can be more immediate, individualized, and open to negotiation (Rollinson, 2005) than teacher feedback, thus providing students with wider perspectives concerning the topic. Such negotiation allows students to “look at the writing from the reviewer’s standpoint” and to “consider the audience for their writing” (Vorobel & Kim, 2014, p. 716). The third advantage is that peer feedback “might increase a range of social and communication skills, including negotiation skills and diplomacy, verbal communication skills, giving and accepting criticism, justifying one’s position and assessing suggestions objectively” (Topping et al., 2000, p. 151). Mendonça and Johnson (1994) investigated 12 ESL students’ peer review process by examining how they negotiated with their peers their intended meaning versus the perceived meaning in the feedback. They found these highly demanding interactions strengthened students’ communicative power. These negotiations

also fostered “a myriad of communicative behaviors” (Villamil & de Guerrero, 1996, p. 69) and improved their diplomacy skills (Hansen & Liu, 2005). In a most recent study, Chang (2016) found Chinese EFL college students (N=13) in Taiwan automatically employed text-based emoticons in their feedback to express positive and negative comments, to mark friendship or to reduce task formality. Therefore, peer review potentially develops a wide range of skills related to giving and accepting criticism, which seems to be indispensable in enhancing writing performance.

Despite these auspicious benefits, disadvantages due to social and cultural domains are also observed. First, because of social or cultural reasons, in non-western cultures in particular, students may also feel uncomfortable or are easily biased in giving feedback (Hyland & Hyland, 2006; Liu & Hanson, 2002; Topping et al., 2000). Second, students’ lack of effective collaborative skills discourages their interaction with each other as reviewers and writers (Lockhart & Ng, 1995; Min, 2008). Third, students from teacher-centered cultures see the teacher as the only source of authority and demonstrate strong reliance on teacher’s feedback (Ren & Hu, 2012; Zhao, 2010).

Linguistic domain. The last domain is specifically correlated with linguistic effects of peer feedback in ESL/EFL contexts. A variety of empirical studies have demonstrated that it creates authentic and interactive contexts for students to discuss language and writing issues, which ultimately leads to improvement of multiple linguistic skills. For instance, Diab’s quasi-experimental study (2011) found that peer feedback assisted

Lebanese university students (N=40) to significantly reduce language errors, thus improving their writing performance. Vorobel and Kim (2014) claimed that oral negotiations in peer response sessions carried “multifaceted benefits for L2 students” (p. 715), including promotion of their speaking and listening proficiency. These findings confirm that the use of L2 in peer review allows students to “consolidate and reorganize knowledge of the L2 in structural and rhetorical aspects and to make this knowledge explicit for each other’s benefit” (De Guerrero & Villamil, 2000, p. 65), and therefore increases their language awareness (Tang & Tithcott, 1999). Together, the findings suggest that “the collaborative dialogue in which peers engage as they work together on writing, speaking, listening and reading activities mediates second language learning” (Swain, Brooks, & Tocalli-Beller, 2002, p. 181).

Some disadvantages of using L2 in peer review are the students’ limited knowledge of the target language and its rhetorical conventions (Hu, 2005; Hu & Lam, 2010). As second language learners, students are still at the stage of developing important linguistic skills, so they may have difficulty in expressing their opinions or giving advice for linguistic reasons (Hansen & Liu, 2005) and identifying language errors when reviewing peers’ writing (Hu, 2005; Hu & Lam, 2010). This might be the reason behind the following statement: “Peer feedback, however, had no significant effect on fluency for either the high- or low-proficient students [N=24]” (Kamimura, 2006, p. 32).

Summary. Peer feedback has been widely used in ESL/EFL writing contexts in past decades and yielded diverse and mixed effects on students' cognition, meta-cognition, affect, social-cultural interaction, and language learning. Peer review activities aid in reducing students' apprehension and increasing their confidence and motivation in writing; they cultivate students' overall writing performance, their advanced higher-order cognitive ability, writer autonomy, and audience awareness. These activities create a supportive environment, improve students' social and communicative skills, and lead to a sense of wider audience; they also enhance L2 students' multiple linguistic skills. However, negative effects have also been found in peer review activities, such as students' discomfort and negativity towards peer review because they lack confidence, social and communicative skills, and the relevant language knowledge.

Given that these effects, whether positive or negative, are potential, careful and systematic implementation is highly recommended as "peer-evaluation groups can be very productive", but "the productivity does not come without a considerable investment of time and effort in preparing students for group work" (Stanley, 1992, p. 230). The following section reviews the empirical studies on implementation of effective peer training and quality of feedback.

Peer Training and Quality of Feedback

Despite the potential negative effects peer review may have on ESL/EFL students, many researchers state that most of these can be offset with sufficient training (Hansen &

Liu, 2005; Lam, 2010; Rollinson, 2005; Topping, 2009). Proper training helps equip peers – affectively, cognitively, socially, and linguistically – to produce more specific and constructive feedback, which nurtures enhanced skills and performance (Berg, 1999; Hu, 2005; Liou, & Peng, 2009; Min, 2005, 2006; Rahimi, 2013; Zhao, 2014). Since the current study incorporates peer feedback into writing instruction in a Chinese university, this section focuses on a literature review of training studies conducted in higher education ESL and EFL writing contexts.

In an early study, Moore (1986) noticed the multiple types of needs for productive peer review and employed a four-part pep talk to assist her students to be affectively ready:

1) You are capable of critiquing others' essays; 2) it's your responsibility to give and take criticism well, remembering that the writers are always ultimately responsible for their own writing, not the evaluators; 3) don't forget to give positive comments; 4) critiquing others' work is useful for you, too—you'll learn skills that will enable you to better evaluate your own work. (p. 23)

To cognitively prepare her students, she then used written guidelines focusing on content and organization of the writings, which refined the students' perspective of good writing. Moore noted that, after training, students became increasingly adept at producing constructive feedback and, "by becoming proficient evaluators of others' work, the students are better able to critically, thoroughly, and objectively evaluate their own writing" (Moore, 1986, p. 23).

Stanley (1992) emphasized the importance of paving the way for constructive peer feedback through lengthy and carefully designed training. To promote students' awareness of what counts as good writing, she familiarized her students (approximately 30) with the genre of the student essay by letting them read a series of papers (from the first to the fourth draft) written by previous students. At each draft, she encouraged students to identify informational gaps, pinpoint unclear sections of text, and report their assessment of the writers' composing process. These activities shaped students' understanding of good writing. Additionally, to affectively prepare the students to provide and accept critique of one's work, she organized a class-wide discussion through which all the participants reached a consensus that students failing to provide specific and substantive feedback would be held accountable. The intention was to motivate students to put forth genuine effort to produce candid and thorough feedback. Stanley's findings exposed that the students who received lengthy training were able to generate more concrete and specific feedback than their untrained peers. Meanwhile, they were more confident in responding to suggestions from their assessors.

Zhu's (1995) experimental study utilized small group conferences to develop university freshman students' (144 native and 25 ESL students) cognitive, social, and linguistic skills to produce profitable feedback in composition classes. After gaining knowledge of some basic peer feedback concepts by watching a demonstration video, Zhu scheduled four-person teacher-student conferences, three times throughout the

semester. During the conferences, the teachers encouraged reviewers to critically identify the strengths and weaknesses of the essay and to provide specific suggestions; they also demonstrated correct strategies for rejecting vague suggestions and seeking clarifications from their reviewers. Zhu (1995) reported that small group conferences significantly contributed to both the quantity and quality of peer feedback.

Based on previous class experiences and suggestions from literature on peer review, Berg (1999) designed an elaborate 11-step training procedure to prepare students to give fruitful peer response. Initial persuasion was used to affectively convince students that they were qualified as reviewers. Response sheets, revision guidelines, collaborative discussion, and videos of successful and unsuccessful peer review were all combined to refine students' (N=46) cognitive, social, and linguistic skills in their reviewing and revising process. The findings showed that the trained group made significantly more content changes and produced significantly better immediate revisions than the untrained group, regardless of the students' L2 language proficiency.

In a 3-year action research study in Singapore, Hu (2005) dedicated himself to discovering effective training approaches to help Chinese ESL university students produce responses which were “clear, specific, supportive and critical” and which “encourage the students to attend to macro-level issues in their own peer responses” (p. 337). In his writing classes, Hu demonstrated a combination of social, cognitive, and linguistic strategies for feedback by “thinking aloud” to model these tactics, such as

identifying both effective and ineffective comments and revisions, pointing out inappropriate language use in responses, discussing various reader stances, and exploring appropriate response behaviors. His modeling led to an increasingly positive attitude toward peer feedback and significant improvement of writing quality concerning both global and local issues.

Noticing the reviewers' misinterpretation of the writers' intentions and the writers' disregard for vague peer comments, Min (2005) focused on cognitive strategies to generate more specific responses on global issues. She used four kinds of evidence-based comments – clarifying writers' intentions, identifying problems, explaining the nature of problems, and making specific suggestions – as guidelines in her training. In addition, she scheduled a student–teacher conference for each student to reinforce in-class training. The findings showed that trained reviewers (N=18) were able to produce substantially more specific suggestions on global issues.

In a variation on Zhu's (1995) study, which utilized small group teacher-student conferences to train students, Min (2006) scheduled a one-on-one conference with each reviewer to discuss their feedback and to explore ways to refine their comments. Her individualized training addressed specific needs of each student and suggested that improved feedback brought about a striking incorporation rate of peer feedback and a significantly higher percentage of successful revisions after training.

Min (2008) reshaped students' (N=18) stances as reviewers through systematic training. She identified four types of reviewer stances in peer interaction – probing, prescriptive, tutoring, and collaborating – which were discussed and demonstrated in the training procedure. The probing and collaborating stances aimed to clarify writers' intentions and to provide suggestions accordingly and thus benefit student writers most. In contrast, prescriptive and tutoring stances tended to push the reviewers' ideas on the writers, which likely discouraged the writers and led to no responses in their revision. After training, the students' prescriptive stance dropped sharply from 59% to 17%, and their collaborative stance increased from 17% to 29%, showing that systematic training redirected students' perspectives of good writing and improved their social skills needed for effective peer review.

Aware of the limitations of providing students with training only at the initial stage of research (Min, 2006), Zhao (2014) emphasized the importance of dynamic and continuous support in her study conducted in mainland China. She incorporated four types of training across five genres throughout the entire research period (four months). She differentiated training foci with varied genres, tailored the training with students' (N=18) ongoing needs, commented on the appropriateness of students' feedback, and helped to mediate disagreements among peer collaborators. The results showed that the quality of peer feedback substantially improved after continuous training.

Summary. The empirical studies discussed above all examined issues concerning peer review training in ESL and EFL writing instruction, in which researchers utilized effective methods to address the social, cultural, cognitive, and linguistic skills needed for fruitful peer feedback. All of them concluded that training significantly enhanced the quality of peer feedback and thus tremendously improved students' writing performance.

Measuring Quality and Outcomes of Peer Feedback and Back-feedback

The literature reviewed in the above section illuminated a significant positive relationship between quality of peer feedback and students' writing performance, but only Min (2006) broadly measured feedback quality by grading and categorizing students' (N=18) comments into three groups: "check plus, check only, and check minus" (p. 124). Are there specific methods to quantitatively measure peer feedback quality? What are the relationships between the quality of feedback/back-feedback and students' learning outcomes? A number of studies shed light on these issues.

Li, Liu and Steckelberg (2010) investigated the relationship between the quality of peer feedback and the quality of the students' (both receivers and reviewers) WebQuest projects. To measure the quality of peer feedback, they utilized a 10-point rubric based on the students' identification of major issues and the quality of the constructive suggestions they provided. Data analysis from 43 undergraduate participants indicated that when controlling the quality of the initial projects, the quality of peer feedback was a strong predictor of the quality of the reviewers' own final project and the quality of the receivers'

immediate project. However, it was a weak predictor of the quality of the receivers' own final project. In other words, students who were capable of providing more constructive feedback performed better in the long run.

Unlike Min's (2006) broad rating category, Sluijsmans et al. (2002) uniquely developed a scoring rubric including seven variables (use of criteria, positive comments, negative comments, constructive comments, posed questions, naïve word use, and structure) to measure Dutch student teachers' (N=93) assessment skills. Their findings demonstrated that the trained groups outperformed the untrained groups in both the quality of evaluative skills and their final performance in creative lesson design. In other words, peer training greatly promoted the quality of peer feedback; however, the relationship between the quality of peer feedback and performance remained unexplored. Additionally, the focus of the study was not writing instruction.

Based on the work of Sluijsmans et al. (2002), Prins, Sluijsmans and Kirschner (2006) developed a Feedback Quality Index together with a scoring rubric ranging between 0 and 100 to analyze the quality of the feedback reports written by general practitioners and their trainers (N=58). The findings exhibited that when the feedback the participants received was objective and clear, the receivers' writing definitively improved.

Gaining insights from Sluijsmans et al. (2002), Kim (2005), and Prins et. al (2006), Gielen, Peeters, Dochy, Onghena, and Struyven (2010) applied a scoring rubric to

measure Belgian secondary students' (N=43) written feedback in their L1 writing. The quality of their feedback was measured by appropriateness to the assessment criteria, specificity, presence of justifications, presence of suggestions for improvement and clear formulation. Their results suggested that receiving justified feedback could significantly improve receivers' performance, which echoed Prins et al.'s (2006) findings. However, both studies left the relationship between the feedback students produced and their own performance unexamined.

Drawing on the previous work of Prins et al. (2006), Gielen and De Wever (2015) graded participants' (N=168) feedback for a Wiki project by assessing three aspects of their feedback: the use of criteria (content and explanations), the nature of feedback (positive and negative remarks), and writing style of feedback (structure and use of key words). Their results evidenced that peer feedback templates (both basic and elaborated) significantly strengthened the quality of students' feedback and students' product scores over time. Nevertheless, their research focus was the relationship between varying structured peer feedback templates and the students' product performance, but not the relationship between feedback quality and students' performance. Nor was it an experimental study.

Differing from the above studies, Kim (2005) designed two scoring rubrics to measure both the peer feedback and back-feedback quality of 82 students to investigate the relationships between the quality of peer feedback and back-feedback and the

receivers' performance in a conceptual organizer task. The feedback scoring rubric evaluated both objective (minimum nodes, lines, shapes, etc.) and subjective (clarity of structure, completeness, support, and creativity) aspects in the peer feedback. The back-feedback scoring rubric evaluated the quality of peer back-feedback by assessing the writers' agreement and disagreement of the suggestions their peers provided to their writing. Kim's study shed new light on the value of measuring students' revising ability, which is an important aspect of the reviewing process. Again, rather than a study concerning writing instruction, it was focused on conceptual organizer task.

Summary. These studies lent support to quality measuring and revealed insights into the relationships between feedback quality and students' performance. However, no studies to date have been conducted in the EFL writing instruction. Although Min (2006), Gielen et al. (2010), and Gielen and De Wever (2015) have made pleasing attempts in their writing practices, their research foci were peer training effects.

Research Questions

While there is consensus on the effects of teacher feedback, there is very little on peer feedback. Its influence on EFL students' English writing performance needs to be explored. In studies concerning writing instruction specifically, most studies included small samples (less than 100) and were conducted qualitatively. Although Zhu (1995) conducted a quasi-experimental research study with 169 participants, 144 of them were

L1 students and the study was carried out 20 years ago. There is an urgent need for a rigorous quantitative study with a larger sample in an EFL context.

In addition, it is surprising that none of the studies specifically and quantitatively examined the relationship between quality of peer feedback and improvement of the students' English writing performance. Perhaps even more surprising is that all of the studies focused solely on peer feedback and not one study examined the students' written responses to the peer feedback received. How students perceive the suggestions and comments provided to their writing, or what is their back-feedback to the comments received, is a significant gap in the literature.

What are the effects of peer feedback on EFL students' English writing performance?
How does the quality of peer feedback correlate with students' writing performance?
What is the relationship between the quality of the students' back-feedback and their writing performance? This study seeks to address these issues, and the specific research questions and rationale are described below.

Research Question 1

What are the comparative effects of peer feedback and traditional feedback on Chinese university students' English writing performance?

Hypothesis 1. Students in the peer feedback group are able to produce similar or better English writing outcomes compared to students in the traditional feedback group.

Rationale. This study assumes that students in the traditional feedback condition receive, on average, less feedback than students in a teacher feedback condition. Therefore, it is reasonable to presume that the findings of comparative studies are in support of this hypothesis. A number of comparative studies reported that peer feedback yielded similar or even better outcomes in students' writing performance than did teacher feedback (Ruegg, 2015; Topping, 1998); peer feedback brought about slightly more successful revisions on content issues (Yang et al., 2006); peer feedback promoted deeper understanding of the comments received (Zhao, 2010).

Research Question 2

How does the quality of students' feedback/back-feedback influence their writing performance?

Hypothesis 2. The quality of students' feedback can be a powerful predictor of students' English writing performance. The quality of students' back-feedback can also be an important predictor of students' English writing performance.

Rationale. Some studies revealed that the ability to produce specific and substantive feedback has significant predictive power on student reviewers' own performances (Demiraslan Çevik, 2015; Li, Liu, & Steckelberg, 2010; Lundstrom & Baker, 2009). Compared to receiving feedback, evaluating others' assignments forces reviewers to

engage more deeply in peer feedback activities, which fosters better internalization of newly acquired skills and knowledge.

Writers' ability to utilize feedback is also an important consideration. The ability to critically assess and select helpful suggestions from feedback received is significantly related to student receivers' performance (Li, Liu, & Zhou, 2012). Unlike passive acceptance of feedback, decisive choices among suggestions also promote student writers' critical thinking abilities.

Summary

Effective writing pedagogy in higher education has been a consistent goal of researchers and instructors in the ESL and EFL writing field. Feedback, a key factor in pedagogical writing practices, has been receiving growing interest, especially after the introduction of peer feedback in the past 20 years. This careful review of existing literature shows that teacher feedback cannot always generate sufficient improvement of university students' writing due to the teachers' heavy workload, their limited perspective of the nature of feedback, and the low trust level between teachers and students.

Additionally, the existing literature has not facilitated a consensus about the effects of peer feedback. So far, there has not been a rigorous quantitative study done with EFL students, and thus the role peer feedback plays is still unclear. Much remains unknown about how the quality of peer feedback affects students' writing performance. Even less exploration has occurred regarding how the quality of peer back-feedback affects

students' writing performance. To address these gaps, this study adopted a quasi-experimental design with a larger sample.

CHAPTER THREE

METHODOLOGY

Overview

This study aims to ascertain the effect of peer feedback on Chinese university EFL students' English writing performance. This chapter reviews the methods utilized in this study, including descriptions and justifications for their appropriateness for the research design. Specifically, it centers on 1) whether peer review activities can nurture improvement in students' writing performance similar to those achieved by traditional feedback; 2) whether the quality of students' feedback / back-feedback is a strong predictor of their writing performance.

Design of the Study

This study was a purely quantitative one, which employed a quasi-experimental nonequivalent control group design (Fraenkel, Wallen & Hyun, 2015). Eight intact classes were randomly assigned to the control and the experimental groups, with four classes in each group. The control group received the traditional feedback. The experimental group was trained on how to generate and use peer feedback, but did not receive feedback from the instructor. The instructors were randomly assigned to four classes from the pool of eight classes.

This study lasted 15 weeks (see Appendix E): in Week 1, the course syllabus and objectives were introduced and discussed; in Weeks 2 and 15, pre-and post-tests (respectively) were administered to all participants; in Week 13, feedback was collected from the peer feedback group only; in Week 14, back-feedback was collected from the peer feedback group only; from Weeks 3 to 14 (four writing cycles, three weeks per cycle), the traditional feedback group received traditional feedback, whereas the peer feedback group received peer feedback (excluding training and follow-up teacher-student conferences on Weeks 13 and 14).

Setting

This study was conducted in English writing classes with English sophomores at a major university of a Southwest province in China. The province is a mountainous one where there are few foreign capital enterprises and joint capital companies. Each year, there are about 240 students enrolled in the School of Foreign Languages, but most of them do not choose English as their major when applying to the university.

The School of Foreign Languages has three departments whose target language is English: the English department, the Translation department, and the Business English department. In the first year, English majors are not assigned to any department but receive instruction concerning general English learning skills: listening, speaking, grammar, and reading. Students have the option to apply to the Translation and the Business English departments at the beginning of the second year, with each department

recruiting a class with a maximum of 30 students. If more than 30 students apply to each department, the students are chosen based on their first-year academic performance. The remaining students are assigned to one of the six classes in the English department, the sizes of which vary from 20-30, but the classes are not balanced in terms of the students' academic skill level.

English writing, a mandatory two-year course, is offered to English sophomores and juniors once a week for two class periods (100 minutes in total). The educational system of the School of Foreign Languages is primarily examination-oriented, with certain classes focusing on exam practice. During the second academic year, the sophomores are instructed chiefly on how to write a traditional argumentative essay of no less than 200 words within 45 minutes, which is the requirement for the writing portion of the Test for English Majors at their fourth term (TEM-4), a nationwide test for all English sophomores.

The two second-year English writing instructors apply a writing-as-a-process approach to their classes, teaching usually 3-4 classes per term. Although the same textbook is used, individual instructors sometimes add extra teaching materials according to their assessment of the students and any deficiencies in the textbook content. Both instructors provide students with constructive feedback, either oral or written, but due to the large ratio of students to instructors, not every individual student has the opportunity to receive teacher feedback on every assignment.

Participants

Sample Size

The required sample size was determined by the research design and the statistical analysis used. This study used a quasi-experimental between group design including an experimental (peer feedback) and a control group (traditional feedback). In addition, a multiple regression was used including the experimental group only. For the quasi-experimental design, a minimum of 30 participants per group is recommended (Fraenkel, Wallen, & Hyun, 2015, p.103). Since there were two groups in this study, a minimum of 60 students was needed. A total of 198 students (101 in the experimental group and 97 in the control group) from eight intact classes participated in this study, thus satisfying the requirement. Multiple regression analysis requires a sample size of a minimum of 20 subjects per independent variable (Ho, 2013, p. 295). Since there were three independent variables in the multiple regression analysis, at least 60 subjects in the experimental group were needed. Four intact classes containing a total of 101 students served as the peer feedback group, thus meeting the subject requirement.

Sample Characteristics

The participants in the study were all from the School of Foreign Languages: six classes from the English department, one from the Translation department, and one from the Business English department. They are English sophomores aging between 18 and 21. The male-to-female ratio was around 1:8. None of the students had taken an English

writing course in the first academic year, nor had they been previously exposed to peer feedback activities.

Instruments

Feedback Guidance Sheet

On the basis of previous research (Berg, 1999; Min, 2006), I designed a Feedback Guidance Sheet (Appendix B-1). The Feedback Guidance Sheet consists of prompts for qualitative comments on a total of eight aspects of writing: six global, one local, and one summary of the strengths and weaknesses of the whole essay. The heavy focus on global aspects is based on the fact that global aspects are viewed as signs of better writing (Liu & Sadler, 2003; Min, 2006). Additionally, a requirement of no less than 150 words is established because quantity is likely to have a positive correlation with quality by encouraging the students to dig more deeply into their ideas and structure.

Back-Feedback Guidance Sheet

The Back-Feedback Guidance Sheet (Appendix B-2) was used to prompt the writer's response to the peer reviewer. The Back-Feedback Guidance Sheet consists of open-ended questions specifically related to suggestions provided in peer feedback. Additionally, the last item asks for a description of three aspects the writer would revise in the second draft. There is also a word requirement of no less than 150 words for students' back-feedback.

Essay Scoring Rubric

I selected the Scoring Rubric (Appendix C-1) used by Yang et al. (2006) to assess the writing proficiency of the participants in this study, because 1) it was designed to measure the writing performance of Chinese university English majors, who are similar to the participants in the current study; and 2) it assesses the essay in a holistic way. This rubric uses a six-point scale ranging from 1 (severely limited) to 6 (impressive).

Feedback Grading Rubric

Based on the previous research of Kim (2005), Sluijsmans et al. (2002), and Prins et al. (2006), I created the Feedback Grading Rubric (Appendix C-2) to measure the quality of peer feedback. The quality of peer feedback is assessed by the identification of major problems in the eight aspects in the Peer Feedback Guidance Sheet and the quality of constructive suggestions for revision. For each aspect, if a student provided no or irrelevant written comments on the essay, “0 points” were assigned. If a student provided comments, but not clearly suggesting how to revise the draft, “1 point” was assigned. If a student provided comments that clearly and reasonably suggested what and how to revise the second draft, “2 points” were assigned. Additionally, one point would be added if the feedback met the word requirement. Therefore, the score range was 0-17.

Back-Feedback Grading Rubric

Also following Kim (2005), I created the Back-Feedback Grading Rubric (Appendix C-3). The items in the Back-Feedback Guidance Sheet corresponded to those items in the Peer Feedback Guidance Sheet. For each aspect, if a student provided no or irrelevant responses to why he/she accepted or rejected the suggestions from the feedback, “0 points” were assigned. If a student provided responses, but did not clearly explain why he/she accepted or rejected the suggestions from the feedback, “1 point” was assigned. If a student provided responses, and clearly explained why he/she accepted or rejected the suggestions from the feedback, “2 points” were assigned. In addition, one point was added if the feedback met the word requirement. The score range for the Back-Feedback Grading Rubric was also 0-17.

Pre- and Post-Tests

To measure the participants’ L2 writing proficiency, two timed argumentative essays served as pre- and post-tests, which were intentionally created with conditions similar to the writing portion of TEM-4. For both tests, the participants were first required to read and summarize the opinions of pros and cons provided, and then took a position to write an argumentative essay of no less than 200 words within 45 minutes. Both the pre- and post-tests prompts (Appendix D) used in this study were argumentative in nature. The topics in the pre-and post-tests were avoided during practices for the writing sessions in the research.

Task

All students in both the traditional feedback and peer feedback groups produced essays on four different topics in four “writing cycles” (see Appendix F) as suggested by Tsui and Ng (2000). In each writing cycle, all participants wrote an argumentative essay of no less than 200 words on the same current controversial topic, two drafts for each topic. Each writing cycle lasted three weeks and each week consisted of two class periods of 100 minutes.

Writing Task for the Control Group

All students in the traditional feedback group participated as writers only. The writing task for the traditional feedback group was similar to that of the peer feedback group. In the first week of each writing cycle, all students participated in group work and brainstorming. After class, all participants produced their first drafts and submitted them upon the due date. Every second week, they were given two volunteers’ sample essays with substantial teacher-written feedback. They were required to review these comments, compare their writing with these samples, and finish assignments such as writing thesis statements, topic sentences, and revising paragraphs. Every third week, they were involved in activities similar to every second week: revising their first drafts and submitting their second drafts.

Writing Task for the Experimental Group

All students in the peer feedback group participated as both writers and reviewers. During the first week of each writing cycle, all students were involved in group work and brainstorming. After class, all students produced their first drafts and submitted them upon the due date. Every second week, each pair of students used L1 to make oral comments and discuss each other's essays face-to-face, by cell phone, or online chat. They then independently generated written feedback on their partner's paper in English. Before submission of their feedback, they were required to take part in the online asynchronous instructor-students conference, where they were offered a sample of feedback (see Appendix G-2) and they had opportunities to discuss the peer feedback they produced. After the conference, they were given two more days to revise and submit their feedback. The tasks of every third week were similar to those in every second week; the only difference was that revising issues were addressed and a back-feedback sample (see Appendix G-3) was provided in this session. The second drafts and back-feedback of each assignment were due every third week.

Peer Feedback Task for the Experimental Group

This research employed face-to-face paired peer feedback because pair work fosters more profound discussion (Hu, 2005; Paulus, 1999), and face-to-face interaction usually leads to more meaning revision (Ho, 2015; Liu & Sadler, 2003). The partner selection was randomized to avoid potential bias due to friendship (Carson & Nelson, 1996) or

boredom resulting from working with the same peer (Ge, 2011). All participants were required to first make oral comments in Chinese for a larger quantity of comments, a fuller understanding of each other (Kamimura, 2006; Zhao, 2010), and for achieving a narrower focus on content and organization (Yu, & Lee, 2014). Next, they were required to submit typed feedback in English with a purpose to practice their target language, and for convenience of the paired students to review the comments whenever needed. The two-step design was chosen because “the combination of written and oral feedback during peer review sessions is essential for enhancing the quality of peer review” (Vorobel & Kim, 2013, p. 714).

Training Procedure

In order to ensure the validity of the research, prior to the training, I scheduled a conference with the other instructor to negotiate the general objectives of the writing course, which included:

1. The objectives of the course are to develop vocabulary, generate ideas, build up students’ abilities of critical thinking, promote collaborative learning, increase students’ confidence in academic writing, and shape long-term writing habits.
2. Both instructors should use the same designated textbooks and essay scoring rubric.
3. The class lecture should incorporate the essential writing components concerning both global and local issues, featuring argumentative essays throughout the entire term.

Brief interviews with the instructor of the traditional feedback group over the research period were held to ensure that he adhered to what had been done in prior writing classes. Any diversion from traditional feedback was discussed with the traditional feedback group instructor and noted in the analysis.

Training to the Control Group

This group received three-step traditional writing instruction throughout the entire research period. An instructor with more than 10 years of English teaching experience taught the traditional feedback group in the same way he had done in prior composition classes.

Instruction of writing knowledge. Given that knowledge of language structures and different types of genres is the foundation for writing, the instructors armed participants with writing knowledge via continuous instruction throughout the whole semester. In this study, lectures on writing knowledge (first class period, 50 minutes per week) were integrated into class periods throughout the semester, except in Weeks 1, 2, and 15 (see Appendix E).

Teacher demonstration. During the second class period (50 minutes) each week, while the peer feedback group received peer training, the traditional feedback group received traditional writing instruction. Typically, the instructor distributed two volunteers' essays to the whole class to read and discuss briefly, one essay at a time. The

instructor then demonstrated his written comments on these sample essays and explained why these comments and suggestions were provided. He further stressed major issues shared by the entire class by comparing the students' discussion with his feedback.

Follow-up reflection. After class, the traditional feedback group were given another two volunteers' essays with substantial teacher-written feedback. They were required to read these sample papers thoroughly and reflect upon the teachers' written feedback. They had an opportunity to schedule an online or face-to-face conference with their instructor to discuss these essays or their first drafts, but this opportunity was not guaranteed.

Training to the Experimental Group

I, serving as the instructor of the peer feedback group, provided the peer feedback group a five-step training program, involving instruction on writing knowledge, initial persuasion, teacher modeling, group practicing, and follow-up teacher-student conferences.

Instruction of writing knowledge. As literature suggests that cognitive abilities are one of the basic foundations for effective peer review, I provided participants with writing knowledge instruction (first class period, 50 minutes per week) throughout the whole research period, which was identical to that of the traditional feedback group.

Initial persuasion. Affective, social, and cultural issues are some other important influential factors in peer review activities. In view of Chinese students' overreliance on teacher feedback (Zhao, 2010), to prepare participants, I first borrowed ideas from Berg (1999) and Moore (1986) to convince the participants of their abilities as reviewers through in-class demonstration of the positive findings from the literature on peer feedback. I then held a class-wide discussion on both reviewers' and receivers' obligations, on the relationship between harmony, friendship, and criticism, and on the social strategies to give criticism and maintain harmony. Initial persuasion was carried out on Week 2 after the pre-test.

Teacher modeling. Potential cognitive, affective, and linguistic issues in peer feedback activities were addressed by teacher modeling. I "thought aloud" (Hu, 2005) the strategies to produce profitable feedback and back-feedback as well as the tactics to selecting constructive feedback. I also demonstrated appropriate linguistic skills to provide, accept, and/or reject suggestions. In the second week of each writing cycle, I used the Feedback Guidance Sheet to "talk aloud" both the strengths and weaknesses in the essays and wrote down my feedback on the sheet. I then used the Feedback Scoring Rubric to explain how my feedback was scored. In the third week of each writing cycle, I used the Back-Feedback Guidance Sheet to demonstrate how to respond to the feedback received (the same one used for comments in the second week). I also "reflected aloud" on why I accepted or rejected the suggestions, wrote down my back-feedback, and used

the Back-feedback Scoring Rubric to explain how my back-feedback was scored. I provided teacher modeling repetitively because constant and repeated reinforcement of students' understanding of the evaluative criteria helped reshape the students' perceptions of good English writing (Ma, 2010), and because clarity about the scoring rubric aided in lessening students' negative attitudes toward peer feedback (Falchikov, 2005).

Group practice. After teacher modeling, all participants were given the same essay to review because modeling with follow-up practice provided students greater opportunities to reinforce their understanding of what had been demonstrated (Van Steendam, Rijlaarsdam, Sercu, & Van den Bergh, 2010). In this session, the students were randomized into groups of four, with students in each group then randomized into pairs within that group. The pairs took turns to practice reviewing (each second week) or revising (each third week), 15 minutes per pair, and 30 minutes in total. When one pair was practicing, the other pair observed and made comments after their performance.

Follow-up teacher-student conference. I created for each experimental class an online group, where I provided an example of strong peer feedback and back-feedback containing language structures (See Appendix G) useful to the novice reviewers. Language structures need to be provided for novice reviewers as they are still learning the target language and may have difficulty expressing their opinions (Hanson & Liu, 2005). All students were required to log in when they were available to read and discuss each

other's feedback before submitting their own feedback or back-feedback. On Weeks 13 and 14, when the peer feedback and back-feedback assignments were collected, follow-up teacher-student conferences were stopped to avoid any potential influences on individual work from the instructor and other participants.

Data Collection Procedure

Data

Pre-test. A pretest was conducted to assess the initial writing ability of the participants to inform the answer to Research Question 1. The topic of the writing prompt was "Should we revive traditional Chinese characters or continue using simplified characters?"

Post-test. The post-test essays were also collected as data for assessing and comparing the writing quality of the two groups to answer Research Question 1. Another argumentative essay "Could online courses be an effective alternative for the traditional in-class course in university education?" was given to all the participants to write.

Feedback and back-feedback assignments. In Weeks 13 and 14, the participants' feedback and back-feedback assignments were collected, respectively, for data analysis. Both in-class training and follow-up teacher-student conferences were stopped in both weeks to ensure that students' feedback and back-feedback assignments were not influenced by the instructor.

Ratings

To ensure the reliability of rating, three instructors in the School of Foreign Languages (including me and two additional teachers) served as raters of the pre-test and post-test essays (Koo & Li, 2016). In chapter Four, I report inter-rater reliability for each pair of raters (1-2, 1-3, 2-3) by computing a Pearson Correlation Coefficient. To satisfy the requirement of an overlap of 15-20% sample for the reliability of the pre-test, I randomly selected 40 essays from the total of 198 pre-test essays to be graded by every rater. The average score of the three raters for each student was used to represent the 40 participants' writing performance in the pre-test. I then divided the remaining 158 essays equally among the raters (about 52 per rater). In a total, each rater graded about 92 pre-test essays, with 40 essays receiving scores from all raters. As the pre- and post-test papers were graded within a relatively short time frame, the inter-rater reliability of the pre-test also applied to the post-test. For the post-test, all 198 test papers were equally distributed to the three raters, 66 for each. Each rater graded a total of about 158 papers, including both the pre- and post-tests. Moreover, I scored the Feedback and Back-Feedback materials alone, because I was most familiar with the topic and instruments.

Data Analyses

Variables

The independent variables included the intervention (traditional feedback or peer feedback), L2 proficiency, feedback scores, and back-feedback scores. There were two dependent variables: pre-test and post-test scores (writing performance).

L2 Proficiency Scores. L2 proficiency in this study was measured by a sample Test for English Majors (TEM). TEM was established by the Chinese State Education Commission in 1991, and has been organized by the national Foreign Language Teaching Committee of Higher Education since then.

There are two levels in TEM, TEM-4 and TEM-8. TEM-4 is for English majors nationwide when they are in their fourth term. TEM-4 aims to assess students' integrative English language ability, including listening, language knowledge, reading, and writing. An authentic test chosen from previous TEM-4 tests was given to all participants immediately before the peer feedback intervention. The scores of the sample test can provide solid evidence of the students' L2 proficiency.

Feedback and Back-Feedback Scores. The scores of the participants' feedback and back-feedback were used as predictors of participants' writing performance. Following Saito and Fujita (2004), I considered the peer feedback and back-feedback of the first three assignments as extra practices for students to familiarize themselves with peer

feedback activities. Thus, I only graded the peer feedback and back-feedback of the fourth assignment, and used the scores as predictors of participants' writing performance.

Initial Writing Ability. Pre-test scores were viewed as a reflection of participants' initial writing proficiency and were used to represent their initial writing ability.

Writing Performance Scores. Pre/post-test scores were regarded as a reflection of participants' initial /latest writing proficiency and were used to represent participants' initial /final writing performance. Both pre- and post-test essays were assessed in a holistic way, using a six-point scale rubric ranging from 1 (severely limited) to 6 (impressive).

Layout of the Analyses

As a preliminary analysis, an independent T-test was carried out to compare pre-test essay means between the traditional feedback and peer feedback groups. The pre-existing differences were reported to help analyze the group difference in the post-test across sections of the writing class. The layouts of the main analyses were summarized in Table 3.1 and Table 3.2.

Table 3.1 Variables for ANCOVA

Quasi-experimental nonequivalent control group design (control + experimental groups)	
Independent variable	● Intervention (traditional feedback vs. peer feedback)
Covariate	● L2 proficiency
Dependent variable	● Writing performance (post-test scores)

The current study applied ANCOVA analysis to compare the effects of peer feedback to traditional feedback. I selected ANCOVA analysis because, it allowed me to examine whether there was a mean difference in the students' post-test performance scores based on the type of intervention utilized with certain influential predictors as covariates. In addition, if pre-existing difference was found, it allowed me to make pre-post comparisons between the traditional feedback group and the peer feedback group while controlling for impact of the students' L2 proficiency.

Table 3.2 Variables for Multiple Regression

Correlational analysis (experimental group only)	
Independent variables	● Initial writing ability (pre-test scores; predictor) ● Feedback scores (predictor) ● Back-feedback scores (predictor)
Dependent variable	● Writing performance (post-test scores)

Following Li, Liu, and Steckelberg (2010), this study conducted a three-step hierarchical multiple regression analysis. I selected a multiple regression because it allowed me to analyze the prediction of each independent variable on the dependent variable while controlling for impact of independent variables entered at later stages. More importantly, it allowed me to specify the order of entry of the independent variables on the basis of logical or theoretical reasons (Ho, 2013, p. 294). The hierarchical multiple regression analysis was conducted with a new predictor added at each step. To investigate the specific prediction of feedback and back-feedback, the initial writing ability variable (pre-test scores) was entered first, because students' previous writing ability largely determines their writing outcomes and should be controlled to examine the specific predictive power of the quality of feedback and back-feedback. Then feedback scores were added on to examine their prediction on students' writing performance, as literature suggests that providing feedback is more effective in enhancing students' writing outcomes than receiving feedback (Demiraslan Çevik, 2015; Li, Liu, & Steckelberg, 2010; Lundstrom & Baker, 2009). Back-feedback scores were added at the last step to investigate their predictive power on students' writing performance. Currently, their predictive power is not clear since little literature has examined it. For all tests, the alpha level was set at .05.

Summary

There is an increasing need for more current research that quantitatively inspects primary pedagogical effects on university students' writing performance in EFL contexts. The purpose of this study is to examine the effect of peer feedback on Chinese EFL university students' writing performance. It also aims to closely investigate the relationship between quality of students' feedback/back-feedback and their writing performance.

This chapter summarizes the methodology used in this study. It describes the quantitative design (a 15-week quasi-experimental control group design), the types of analysis (ANCOVA and hierarchical multiple regression), the settings (Chinese EFL context), and the entire procedure.

CHAPTER FOUR

RESULTS

This quantitative study investigated the effects of peer feedback instruction on Chinese English sophomores' writing performance by, 1) comparing the effects peer feedback and traditional feedback had on students' writing achievement; and 2) examining the relationship between writing performance and the quality of students' feedback/back-feedback. An independent T-test (preliminary analysis), a one-way ANCOVA, a one-way repeated measures ANCOVA together with a series of multiple regression analyses were employed to analyze the data collected for this study.

The following two research questions (RQ) were addressed and have been presented in this chapter:

RQ1. What are the comparative effects of peer feedback and traditional feedback on Chinese university students' English writing performance?

RQ2. How does the quality of students' feedback/back-feedback influence their writing performance?

For the purpose of presenting the results, Chapter Four reported preliminary data analysis of each test first to ensure that the assumptions for parametric statistics were upheld. Next, the results of every test were reported. An alpha level of .05 was used for all statistical tests.

Preliminary Data Analysis

Interrater Reliability

As a Pearson Correlation Coefficient value higher than .70 indicates an acceptable interrater agreement (Barrett, 2001), the r values between each pair of raters [$r=.69$ (raters 1-2), $r=.73$ (raters 1-3), $r=.76$ (raters 2-3)] in this study showed a generally reasonable interrater reliability. In addition, Intraclass Correlation Coefficient (ICC) was used to test the interrater reliability among the three raters. According to Koo and Li (2016), an ICC value between 0.75 and 0.9 suggests reasonable reliability. Thus, the ICC result ($r=.89$) satisfied this criteria.

Comparing Group Difference with No Covariate

Before answering the first research question, I conducted a preliminary independent T-test to determine whether the two groups began with similar writing ability. If the groups were found to differ in their initial writing ability, the difference would be taken into consideration in the subsequent ANCOVA analysis. Writing scores were based on a scale of 1-6 for both the pre- and post-tests.

Testing T-test Assumptions. Prior to running the independent T-test, the assumptions of normality and homogeneity (Ho, 2013) were examined.

Normality. Given that the T-test and ANOVA findings were quite robust regarding departures from normality (Posten, 1984; Schmider, Ziegler, Danay, Beyer, & Bühner,

2010), the results can be accurate even when the Skewness was 2 and Kurtosis was 9. In this study, the Skewness was .61 and Kurtosis was -.08 for the peer feedback group, and the Skewness was 1.04 and Kurtosis was .85 for the traditional feedback group; hence, all of the metrics were smaller than 2 or 9. Therefore, an independent T-test was appropriate to examine the difference between the two groups though the data was non-normalized.

Homogeneity. The assumption of homogeneity was examined by the Levene's Test for Equality of Variances. The result showed that there was no difference in the two observed sample groups [$F=.005, p=.946$], thus homogeneity was accepted.

Results of the Independent T-Test analysis. The results revealed a marginal difference in English writing ability between the traditional feedback and peer feedback groups ($t_{196} = 1.99, p = .048$). To better understand the difference, the results are shown in Table 4.1, in which the mean and standard deviation for the traditional feedback and peer feedback groups are shown in the first and second column, respectively, and the relevant T-test value and its significance are shown in the third column.

The results in Table 4.1 shows that the pre-treatment writing performance scores of the traditional group were slightly lower than that of the peer feedback group, with a mean difference of .23 points. According to Cohen (1992), a d value between 0 - 0.2 indicates a small effect size, a d value between 0.2- 0.5 indicates a medium effect size, and a d value beyond 0.5 indicates a large effect size. In this study, the pre-treatment

Table 4.1 Comparing Means Between the Traditional Feedback and Peer Feedback Groups Before Intervention

Traditional feedback Mean (SD)	Peer feedback Mean (SD)	T-test value (<i>p</i> value)
2.86 (.816)	3.09 (.830)	1.99 (.048)

writing performance score difference revealed a marginal medium effect size (Cohen $d = .28$). Thus, when comparing the overall performance between groups without taking English proficiency into account, the students in the peer feedback group scored somewhat higher in their writing, suggesting that this group may have had a slight advantage prior to the intervention.

RQ1: Comparing Instructional Approaches After Covarying English Proficiency

The first research question centered on whether there was a statistically significant difference in the English sophomores' writing performance based on type of feedback (traditional feedback vs peer feedback), controlling for English proficiency (L2). The hypothesis for research question one was that students in the peer feedback group would be able to produce similar or better English writing outcomes when compared to students in the traditional feedback group. After the preliminary analysis of students' pre-treatment writing performance, a one-way ANCOVA was used to test whether the

instructional approaches had any effects on Chinese university students' post-treatment writing performance after taking English proficiency into account.

Testing ANCOVA Assumptions

Before running a one-way ANCOVA analysis, assumptions of normality, homogeneity, linearity, and homogeneity of regression slopes (Keselman et al., 1998) were checked.

Normality. The Kolmogorov-Smirnov test showed that the Skewness was .21 and Kurtosis was .98 for the peer feedback group, and the Skewness was .54 and Kurtosis was 1.44 for the traditional feedback group. As previously discussed, these values would not violate the assumptions in a way that would negatively impact the ANCOVA results.

Homogeneity. The Levene's test showed that there was a difference in the two observed sample groups ($F=23.58, p=.000$), thus the assumption of homogeneity of variances between groups was violated. However, T-test, ANOVA, and ANCOVA are robust with respect to potential violations, any sized difference in the variances would not be a problem when the equal sample sizes are greater than 15 (Ramsey, 1980; Posten, 1984). As this study included 198 participants, the violation of homogeneity did not cause any problem.

Appropriateness of Covariate. The appropriateness of covariate was checked by the between-subjects effect tests. The result showed that there was no statistically significant

difference between the traditional feedback and peer feedback groups measured by English proficiency (L2) ($F=1.99, p=.16$), indicating that the values of the covariate did not vary across the two groups. Thus, the variable L2 was appropriate as a covariate.

Linearity. Linearity was checked using simple regression. The results showed that there was a linear relationship between the covariate and the dependent variable ($t=6.258, p=.000$), thus this assumption was accepted.

Homogeneity of Regression Slopes. The assumption of homogeneity of regression slopes was tested by the between-subjects effect tests, which revealed that the slopes for the covariate (L2) and the independent variable (groups) were similar, indicating that there was no interaction between L2 and Groups as the value for Groups* L2 was ($F=2.08, p=.15$). Thus, the assumption was accepted.

Comparing Instructional Approaches after Taking English Proficiency into Account

The results of the ANCOVA analysis are shown in Table 4.2, in which the adjusted mean and standard error for traditional feedback and peer feedback groups are shown in the first and second column, respectively, and the relevant F -test value and its significance are shown in the third column.

As can be observed in Table 4.2, the results revealed a significant effect of feedback approaches on students' writing performance scores in the post-test after controlling for English proficiency. While not shown in the table, English proficiency (L2) also had a

Table 4.2 Comparing Means Between The Traditional Feedback And Peer Feedback Groups After Taking L2 Into Account

Traditional feedback Adjusted Mean (Standard Error)	Peer feedback Adjusted Mean (Standard Error)	F-test value (<i>p</i> value)
3.09	3.91	69.12
(.070)	(.069)	(.000)

significant impact on students' writing performance [$F(1, 195) = 40.90, p = .000$]. This analysis demonstrated that there were significant instructional group differences favoring peer feedback over traditional feedback after taking English proficiency into account.

Comparing the Impact of Instructional Approaches on the Change in Writing Performance after Taking English Proficiency into Account

As the independent T-test showed marginal initial writing difference between the peer feedback and traditional feedback groups, a one-way repeated measures ANCOVA analysis was conducted to investigate whether the group difference found in the ANCOVA analysis remained after controlling for English proficiency (L2) and taking the initial writing difference into account.

Table 4. 3 shows the results of this two-group repeated measures ANCOVA test. The table consists of two sections: the first section shows the results of the tests for within-subjects contrasts and the second section shows the results for the tests for between-subjects effects. For each section, the first column shows the factors which may

have had impact on writing performance, the second shows the sum of squares, the third shows the degrees of freedom for that factor, the fourth shows the F value, and the final column shows the significance related to that test value.

The within-subjects results showed that the change in writing performance scores differed significantly by type of intervention ($p < .001$) but not by L2 ($p = .62$). In the between-subjects effects, both intervention and English proficiency were significantly related to the writing performance scores when pooled across both time points. In other words, after controlling for English proficiency (L2) and taking the initial difference into account, the change in writing performance scores from pre-to post-assessment was significantly different between the peer feedback and traditional feedback groups.

To better understand the intervention difference in writing performance over time, the adjusted means from this analysis are shown in Figure 4.1. In this figure, each line shows the adjusted mean score in the pre-test and then in the post-test, with the peer feedback group's scores shown in the solid line and the traditional feedback groups' scores showed in the dashed line.

As can be seen in Figure 4.1, the peer feedback group had an advantage in writing performance over the traditional feedback group that increased from an adjusted mean difference of 0.18 points to one of 0.82 points. The peer feedback group made more

Table 4.3 Comparing Means Between The Traditional Feedback and Peer Feedback Groups After Taking L2 and Initial Writing Difference Into Account

Within subjects contrasts

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	.444	1	.444	.940	.334
Time by L2	.116	1	.116	.246	.620
Time by Intervention	9.986	1	9.986	21.113	.000
Error (time)	92.226	195	.473		

Between subjects effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	43.361	1	43.361	72.054	.000
L2	34.460	1	34.460	57.263	.000
Intervention	24.160	1	24.160	40.147	.000
Error	117.349	195	.602		

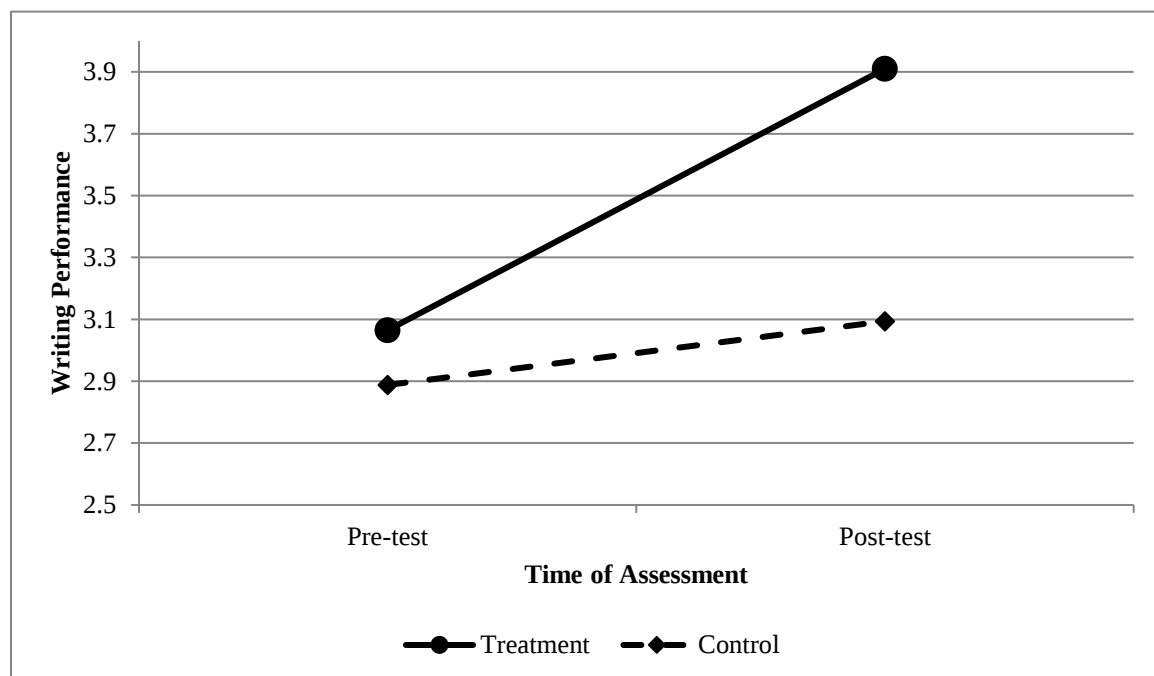


Figure 4.1. Comparing the Change Over Time in Students' Writing from Pre- to Post-treatment between Treatment and Control Groups

improvement than the traditional feedback group, with positive changes of about .85 and .21 points respectively, which demonstrated a very large effect size (Cohen $d=1.17$). These results support hypothesis one that students in the peer feedback group would be able to produce similar or better English writing outcomes when compared to students in the traditional feedback group.

RQ2: Relationship between the Quality of Feedback / Back-Feedback and Writing Performance

Research question 2 addressed the relationships between the quality of feedback/back-feedback and writing performance. The hypothesis for question two

consisted of two sub-hypotheses: 1) the quality of students' feedback can be a powerful predictor of students' English writing performance; and 2) the quality of students' back-feedback can also be an important predictor of students' English writing performance. Both hypotheses were tested using a series of multiple regression analyses.

In Chapter Three, I selected a hierarchical multiple regression to examine whether the quality of feedback and the quality of back-feedback were significant predictors of the students' writing performance. In this design, the order of entry of the variables was: students' initial writing ability first, the quality of feedback second, and the quality of back-feedback third. The sequence was based on the presumption that students' initial writing ability was the most influential predictor and should be entered first to investigate whether the addition of the other two predictors can increase the predictive capacity of the models. However, the results showed that students' initial writing ability was not a significant predictor in any of the three models. One possible factor for this interesting finding might be that the impacts of feedback and/or back-feedback were so strong that the effect of initial writing ability was offset. Given this condition, research question 2 was broken into three sub-questions and a revised design including a series of multiple regression analyses was used to explore how the specific impacts of the two key predictors (feedback and back-feedback) on students' writing performance changed in the models, with the initial writing ability functioning as a covariate. Research question 2 and the three sub-questions are:

RQ2. How does the quality of students' feedback/back-feedback influence their writing performance?

2a When the students' initial writing ability is controlled, what is the relationship between the quality of students' feedback/back-feedback and their writing performance?

2b How does the relationship between the quality of students' feedback and their final writing performance change, after taking their initial writing ability and the quality of their back-feedback into account?

2c How does the relationship between the quality of students' back-feedback and their final writing performance change, after taking their initial writing ability and the quality of their feedback into account?

For the three sub-questions, I used a multiple regression analysis to examine the general predictive capacity of the quality of students' feedback/back-feedback to answer question 2a, and two hierarchical multiple regression analyses to explore how the prediction of the two key predictors changed to answer questions 2b and 2c. In the multiple regression analysis, both the two predictors and the covariate were entered at the same time. In the hierarchical multiple regression analysis for RQ2b, the quality of feedback was entered first, then initial writing ability was added, finally back-feedback was introduced. In the hierarchical multiple regression analysis for RQ2c, the quality of back-feedback was entered first, then initial writing ability was added, finally feedback was introduced.

Testing Multiple Regression Assumptions

Prior to conducting this series of multiple regression analyses, the assumptions of multivariate normality, independence of error terms, linearity, homoscedasticity, multicollinearity, and outlier (Ho, 2013) were assessed through the regression analysis.

Normality. The normal P-P plot of standardized residuals (See Appendix H-1) showed that the points cluster tightly along the diagonal line, which indicated a relatively normal distribution for the dependent variable (Post-test scores).

Linearity and Homoscedasticity. Linearity and homoscedasticity were examined by the scatterplot of standardized residuals against the standardized predicted values (See Appendix H-2). It showed a roughly rectangular distribution with most of the scores clustered in the center, indicating that these assumptions were met.

Independence of Error Terms. The Durbin-Watson d statistic was used to test the assumption of independence of error terms. As the Durbin-Watson d value ($d = 1.69$) was between the two critical values of $1.5 < d < 2.5$, it can be assumed that each predicted value was independent.

Multicollinearity. Multicollinearity was evaluated by the Tolerance and Variance Inflation Factor (VIF) values. Tolerance values higher than 0.10 and VIFs smaller than 10 suggest that there was not a strong overlap between the independent variables. In this

study, the tolerance values for the quality of feedback, for the quality of back-feedback, and for the initial writing ability were .876, .886, and .981, respectively; the VIF values for the quality of feedback, for the quality of back-feedback, and for the initial writing ability were 1.141, 1.129, and 1.019, respectively. All the values met the assumption, indicating that no independent variables were highly correlated.

Outlier. To test whether there were outliers in the data, a Mahalanobis distance test was employed whose maximum value (*Mahal*=11.034) was much lower than the critical value (13.82) associated with two predictors, indicating there were no multivariate outliers (Leys, Klein, Dominicy & Ley, 2018).

Unique Impact of Feedback/Back-Feedback on Students' Writing Performance

The results of the multiple regression testing the effect of the quality of students' feedback and back-feedback on students' writing performance are shown in Table 4.4. As shown in Table 4.4, both of the two key predictors had significant and unique effects on the students' writing performance. The quality of students' feedback had a positive impact on the students' writing performance in the post-test even after the students' initial writing ability and the quality of students' back-feedback were taken into

Table 4.4 Effect of the Quality of Students' Feedback/Back-feedback on Students' Writing Performance, Controlling for Initial Writing Ability

Variable	Standardized Coefficient	<i>t</i> -value	<i>p</i> value
Initial writing ability	.17	1.80	.074
Quality of feedback	.21	2.13	.035
Quality of back-feedback	.24	2.39	.019

$$R^2 = .17, F_{(3,97)} = 6.43$$

account. In addition, the students' back-feedback had a positive impact on the students' writing performance even after the students' initial writing ability and the quality of students' feedback were taken into account. Between the two key predictors, the quality of students' back-feedback had a slightly larger impact, with a standardized coefficient 0.03 higher than that of the quality of students' feedback. These results supported that both the quality of feedback and the quality of back-feedback would have a unique positive impact on students' writing performance.

Examining the Change in the Impact of Feedback

The results of the hierarchical regression examining the change in the impact of the quality of students' feedback on their writing performance after controlling for initial writing ability (Model 2) and then the quality of students' back-feedback (Model 3) are shown in Table 4.5. For those estimates that had a significance test, asterisks indicated the

Table 4.5 Change in the Effect of the Quality of Students' Feedback on Writing Performance, Overall and then Controlling for Initial Writing Ability and the Quality of Students' Back-feedback

Model: Predictors Included	Std. Coefficient (β) Quality of feedback	% Change in β by Model	R^2	Change in R^2 by Model
1: Quality of feedback	.31**	---	.094**	---
2: Quality of feedback Initial writing ability	.29**	-6%	.117**	.023
3: Quality of feedback Initial writing ability Quality of back-feedback	.21*	-28%	.166**	.049*

* $p < .05$; ** $p < .01$

significance of the estimate.

As is shown in Table 4.5, the initial positive effect of the quality of students' feedback on their writing performance was slightly reduced (6%) after controlling for initial writing ability. This positive relationship was then reduced another 28% after taking the quality of students' back-feedback into account. However, there still remained a small but significant impact on students' writing achievement uniquely associated with the quality of students' feedback. This significant relationship suggested that the quality of students' feedback had a stable positive effect on their writing achievement.

Examining the Change in the Impact of Back-Feedback

The final analysis of this chapter addressed a parallel question concerning feedback and back-feedback, examining the change in the impact of back-feedback before and after controlling for the impact of feedback on post-treatment writing performance. The results of the hierarchical regression examining the change in the effect of the quality of students' back-feedback on their writing performance after controlling for initial writing ability (Model 2) and then the quality of students' feedback (Model 3) are shown in Table 4.6.

As is shown in Table 4.6, the initial relationship between the quality of students' back-feedback and their writing performance actually increased slightly (3%) after controlling for their initial writing ability. This positive relationship was then reduced 23 % after taking the quality of students' feedback into account. However, there remained a moderate and significant impact on students' writing achievement uniquely associated with the quality of students' back-feedback. This result indicated that the quality of students' back-feedback had a positive and stable effect on their writing achievement. The results for both hierarchical multiple regression analyses indicated a support for the effectiveness of the peer feedback approach implemented in the experimental group.

Table 4.6 Change in the Impact of the Quality of Students' Back-feedback on Writing Performance, Overall and then Controlling for Initial Writing Ability and the Quality of Students' Feedback

Model: Predictors Included	Std. Coefficient (β) Quality of feedback	% Change in β by Model	R ²	Change in R ² by Model
1: Quality of back-feedback	.30**	---	.089**	---
2: Quality of back-feedback Initial writing ability	.31**	+3%	.127**	.038*
3: Quality of back-feedback Initial writing ability Quality of feedback	.24*	-23%	.166**	.039*

* $p < .05$; ** $p < .01$

Summary

This chapter tested: hypothesis 1) students in the peer feedback group would be able to produce similar or better English writing outcomes when compared to students in the traditional feedback group; and hypothesis 2) the quality of students' feedback and the quality of students' back-feedback can be powerful predictors of students' English writing performance. The results of the independent T-test and one-way repeated measures ANCOVA revealed that beginning with slightly different writing ability, the peer feedback group made more growth in writing achievement than the traditional feedback group in the post-test. Further analysis through hierarchical multiple regression analyses showed that both the quality of students' feedback and the quality of students'

back-feedback were significant predictors of students' writing performance, when the prediction of the independent variables entered at later steps were controlled for. In comparison, the quality of students' back-feedback had a slightly stronger predictive capacity than the quality of students' feedback. The results of all these tests supported both hypotheses.

CHAPTER FIVE

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

The purpose of this study is to ascertain the effect of peer feedback on Chinese English sophomores' writing performance. Chapter Five first summarizes and interprets the results of the tests. Next, the implications and limitations of the study are discussed. Finally, recommendations for pedagogical practices and future research along with a brief conclusion are presented.

Conclusions and Discussions

The Effects of Peer Feedback on Chinese English Sophomores' Writing Performance

The one-way repeated measures ANCOVA results showed that the peer feedback group outperformed the traditional feedback group significantly, even after the students' English proficiency and their initial writing ability were taken into account. This is in line with Simpson and Clifton (2016) and Topping (1998; 2009) that peer feedback can yield an overall similar and even better results on students' writing performance.

Extensive training assisted the participants in the peer feedback group to better understand the criteria for good writing. After repeated peer review practices, the participants became skilled in identifying problems in their peers' essays. More importantly, as each student only commented on one partner's essay at a time, they had

sufficient time to discuss and negotiate wider perspectives concerning the topic, which nurtured fuller understanding (Zhao, 2010) as well as richer, more immediate, and individualized responses (Rollinson, 2005). Consequently, true growth in writing performance was realized. This finding is consistent with many studies reporting the effectiveness of peer feedback on L2 writing performance (Diab, 2011; Ho, 2015; Lundstrom & Baker, 2009; Ruegg, 2015; Yang, 2006; Zhao, 2010, 2014).

In comparison, a large student to teacher ratio reduced the opportunity for the students in the traditional feedback group to receive substantive and constructive feedback. Limited amount of either oral or written feedback from the teacher on assignments may negatively impact their writing gains.

The Relationships between Quality of Peer Feedback/Back-Feedback and Writing Performance

Feedback contribution. In this research, the quality of students' feedback was a stable predictor of students' writing performance in all the three hierarchical multiple regression models. Even after students' initial writing ability was controlled for in model 2, and the quality of back-feedback was taken into account in model 3, it remained a significant predictor. This supported the hypothesis that the quality of feedback students provided to their peers was strongly and positively related to their writing performance. This finding is in line with some current studies (Demiraslan Çevik, 2015; Li, Liu, & Steckelberg, 2010; Lundstrom & Baker, 2009) that show providing constructive feedback

can engage students in an active learning mode, which ultimately sharpens students' reviewing and learning abilities.

One possible reason for the strong correlation between feedback and writing achievement in this research might be that producing profitable feedback engaged students in cognitively challenging activities, such as pinpointing problems, seeking clarification, making critical judgments, providing justification, and proposing solutions (Rollinson, 2005; Topping, 1998, 2009). The combination of oral and written feedback design further increased the assessment load which required intensive involvement of higher order abilities. Repeated practices of such activities (Ma, 2010; Min, 2006; Van Steendam et al., 2010) aided in internalizing students' newly obtained skills and consolidating their writing ability.

Back-feedback contributions. While most of the current peer feedback studies simply examined the relationship between peer feedback and writing performance, this study uniquely investigated how the student authors' responses to the reviewers' feedback, or back-feedback, influenced their own writing performance. The hierarchical multiple regression results suggested that the quality of students' back-feedback was most strongly related to their writing achievement (after controlling for the impacts of both the initial writing ability and the quality of students' feedback), which confirms that students' ability to critically assess and apply the feedback received (Kim, 2005; Li, Liu, & Zhou, 2012) is positively associated with students' writing performance. This finding not only

supports hypothesis 2 but also fills the void of knowledge about writing instruction because the role back-feedback plays in peer review activities in both L1 and L2 writing education has been largely ignored and there were few empirical findings reported.

There are some possible explanations for this result. First, producing quality back-feedback is a highly demanding task which involves a process from reflecting upon output one (first draft) to evaluating input (feedback received plus output one) to generating output two (back-feedback and revised draft). These arduous activities demand and, in turn, facilitate development of advanced learning abilities, such as critical thinking and logical reasoning. If students received comments contradictory to their own writing purposes, they tended to evaluate the peer feedback received, reflect upon their personal writing purposes, explore the topic further, and close the gap by making comparisons. Assessing and revising their products with thoughtful comments engaged them in deeper critical thinking and nurtured stronger logical reasoning through meditating on conflicting ideas.

Second, producing back-feedback may trigger deeper audience awareness. Filling out the back-feedback guidance sheet provided students with an in-depth opportunity to meditate on the gaps between what was conveyed and what was perceived. This would raise their awareness of why and how these gaps appeared. In their future writings, they are likely to take their target audience into consideration to avoid potential gaps.

Third, producing high quality back-feedback may encourage positive “skepticism” toward peers’ feedback, which ultimately cultivates more writer autonomy. In performance, students were warned against blind acceptance of suggestions as the quality of peer feedback often varied. Conscious identification and selection of suggestions were highly recommended during the revision process through filling out the back-feedback guidance sheets. Interestingly, this process may have triggered a stronger positive skeptical attitude and deeper exploration, which facilitated the development of writing autonomy and aided in securing long-term growth in writing achievement. This echoed Yang et al.’s (2006) observation that a lack of trust in peer feedback fosters more successful revision and stronger writer autonomy.

Feedback and back-feedback contribution. Both feedback and back-feedback sharpened students’ self-regulated learning skills which, in turn, made true growth in writing occur. Self-regulated skills, with feedback as their “inherent catalyst” (Butler & Winne, 1995, p. 246), feature deep engagement in cognitive and meta-cognitive tasks and emphasize “autonomy and control by the individual who monitors, directs, and regulates actions toward goals of information acquisition, expanding expertise, and self-improvement” (Paris & Paris, 2001, p. 89). These tasks and activities are typical in both the processes of generating feedback and back-feedback, during which students actively produce and use either external or internal feedback to evaluate their current

learning performance, compare it to the desired goals, and take actions to improve their performance.

When producing feedback, students were fully aware of their specific goals: they directed the evaluation process of detecting problems, negotiating with the paired writers, seeking clarification and further information, and providing external feedback. During the process, it was highly likely that they compared their own work with their partner' and applied what they had learned to assess their own work, in which their ability to produce internal feedback developed through generating external feedback (Nicol, Thomson, & Breslin, 2014). These activities expanded their expertise, contributed to their learner autonomy, and led to their self-improvement.

Specifically, the process of producing back-feedback developed more learner control. Unlike the feedback process during which each pair of students were required to make face-to-face negotiations, the back-feedback process involved little discussion between the paired students. Participants depended solely or mainly on themselves to evaluate their first drafts together with suggestions given and then provide internal feedback accordingly.

These feedback/back-feedback tasks were highly self-regulated, which ultimately furthered students' autonomy and control. Consequently, their need for receiving feedback from their friends or teachers could decrease (Nicol et al. 2014). The findings of

this study consolidate the power of formative peer feedback as a tool to maximize learning.

Implications

The findings of this study suggest three important considerations for instructors and researchers who would like to apply formative peer feedback to their classes. First, novice writers may benefit more through peer collaboration. Second, well-designed feedback procedure is indispensable for yielding fruitful feedback and back-feedback. Third, wide and diversified input pave the way for mature and sophisticated ideas.

Novice Writers Benefit More

This study centered on Chinese EFL college writers who have never been exposed to peer feedback activities. Inexperienced in writing academic English essays and giving feedback, they could be identified as novice reviewers and writers. Nevertheless, the result of this research revealed that the peer feedback group outperformed the traditional feedback group in writing achievement, which provided evidence that even novice student reviewers are able to produce profitable feedback. This finding consolidates that becoming critical readers of peers' writing makes it possible for them to be more alert in their own writing, thus cultivating them into conscious writers (Berg, 1999; Foley, 2013; Nicol et al., 2014). Surprisingly, Lundstrom and Baker (2009) observed that students at "lower proficiency level made more gains than those at higher proficiency levels" (p. 30),

which provides further insight into the potential of peer review activities on novice writers.

Well-designed Feedback Procedure

Well-designed feedback procedure is essential for the effectiveness of students' feedback and back-feedback. To ensure the quality of students' comments, this study applied a two-step (Vorobel & Kim, 2013) feedback design (oral comments in native language first and then typed feedback in the target language). At the first step, oral comments led to a larger quantity of comments in which the native language played a facilitative role in maximizing students' understanding of each other (De Guerrero & Villamil, 2000; Kamimura, 2006; Zhao, 2010). At step two, typed English feedback compelled students to practice their target language and gave them convenience to reflect upon comments and suggestions repeatedly whenever needed (Rollinson, 2005). The combination of oral and written feedback not only ensured the quality of peer feedback and back-feedback through a balanced usage between L1 and L2 in the EFL context, but also offset potential drawbacks L1 may have on students' L2 writing performance.

Wide Input

Writing, an advanced and demanding form of output (Vygostky, 1986), is tightly associated with wide input; without a variety of comprehensible input (Krashen, 1981), a student is less likely to be a strong reviewer and writer. This is explicitly reflected in the

quality of students' feedback and back-feedback. During the research, many students reported that the peer responses they received were generally constructive with respect to aspects such as usage of examples, organization, strengths, and weaknesses. Whereas, the feedback concerning the thesis statement and topic sentences was somewhat superficial, especially when the topics were thought-provoking. Though most peer students were able to technically detect whether the thesis statement and topic sentences were appropriate regarding their location, they often failed to provide in-depth suggestions. A close examination of students' back-feedback also demonstrated a lack of specific responses and revision plans on thesis statement and topic sentences.

For instance, when assessing an essay "Gratitude is the Most Important Quality of An Undergraduate Student", the reviewer pointed out that the topic sentences were vague. However, gratitude is a very abstract concept so she could not reveal the essence of gratitude and provide a specific suggestion. Faced with similar issues, in the author's revised draft, she changed the topic sentence into "Gratitude means that one should be grateful to their lives", which was more unclear. In this case, a lack of in-depth understanding of life from both the reviewer and author might be the root of their vague opinion. Under such conditions, students' writing skills alone are not enough to produce quality feedback/back-feedback; rather, their experiences and general knowledge tremendously affect their responses. Consequently, it is imperative to broaden students' horizons by increasing diversified input and providing more opportunities for them to

unfold their interactions with different peer students to expand each other's ZPD (Wells, 1999).

Limitations

Given that all studies have limitations, this study is of no exception. Due to time constraints, this study used a non-randomized sample and examined only one writing genre in a monotonous learning context. Students with low motivation may also have influenced the effectiveness of the study.

Time Constraints

This study lasted only one semester which was not long enough for novice writers to internalize their evaluative and revision skills. Due to the large class size and low teacher to student ratio, it was very hard for every student to get sufficient time and guidance either during the class period or during the online session to discuss their feedback and back-feedback with me individually.

Additionally, because of time constraints, I did not give the student authors' back-feedback and revised drafts to their paired reviewers; this inevitably resulted in the reviewers' loss of in-depth opportunities to reflect upon their own comments and why these suggestions were accepted and/or rejected, thus influencing the quality of their future feedback.

Non-randomized Sample

Although randomization was implemented in the process of peer training, intact classes were used in this research which led to non-randomized sample. Given that the participants cannot represent EFL students in China, the generalizability of this study was more limited, which may cause a problem in determining the external validity of the current study.

Monotonous Genre and Learning Context

This study, using argumentative essays alone as the research focus, excluded other writing genres, which limited the generalizability of applying the findings to future studies concerning writing instruction. The validity of peer feedback, particularly, back-feedback needs to be verified through further studies covering multiple genres. Additionally, the study was conducted in an EFL context, where all the participants took advantage of their native language to achieve mutual understanding. The applicability and effectiveness of back-feedback in an ESL context where learners speak different first languages remains unexplored.

Students with Low Motivation

As described in Chapter Three, many students in this study did not choose English as their major; they were assigned to the School of Foreign Languages when applying to the university. In their self-introduction in Week One, many of them explicitly expressed that

they viewed English as secondary to their other courses such as law, computer science, and administrative management. Their L2 scores also reflected their low motivation toward learning English, which may have had a direct influence on the relationship between feedback/back-feedback and their writing performance ($r_{\text{feedback}} = .31$, $r_{\text{back-feedback}} = .30$), and consequently, the internal validity and effectiveness of the findings.

Recommendations for Pedagogical Practices and Further Research

This study suggests two valuable strategies for pedagogical practices and a number of research areas for further exploration including replicated studies in EFL/ESL/L1 contexts, in-depth quantitative studies, small scale qualitative studies, and further exploration of how peer feedback influences self-feedback abilities.

Extensive and Continuous Training

Extensive and specific training is highly recommended for pedagogical practices because it makes true growth in writing performance possible (Lam, 2010; Min, 2006; Ma, 2011; Zhao, 2014). According to Nunan (2001), writing is the most demanding language skill. Being a constructive writer requires “a variety of motor, cognitive, and affective skills” (Graham & Harris, 2016). This study provided participants with continuous (Zhao, 2014) peer feedback training, covering writing knowledge and language patterns for cognitive and linguistic needs as well as skills for affective and social-cultural concerns. In practice, it was found that fully prepared training (Berg, 1999;

Hu, 2005; Min, 2006; Rollinson, 2005) offset students' negative attitudes, tremendously reduced worries, and increased confidence toward producing peer feedback and back-feedback.

Well-developed Instruments

Another pedagogical recommendation from this study is that well-developed instruments are crucial to the effectiveness of research. In performance, the participants reported that the carefully designed guidance sheets and rubrics helped them understand evaluation criteria easily and avoid focusing highly on local aspects like grammar and mechanics. Despite the clarity, it was still found that one aspect concerning “off topic or not” should be included at the beginning of the guidance sheets. This missing aspect caused the phenomenon of “seeing trees rather than the woods”: some students failed to make a critical overall judgment even though the essay assessed was off topic from the thesis statement. Unconscious of or ignoring the off-topic problem, they continued to evaluate each aspect separately instead of building connections throughout the essay, which led to less profitable feedback.

This finding confirms the crucial role of well-developed instruments (Kim, 2005; Prins et al., 2006; Sluijsmans et al., 2002), which should vary in accordance with different writing genres and diversified instructional contexts (Zhao, 2014). Specifically, in ESL contexts where students speak different first languages, clarified instruments are of great help to avoid confusion.

Replicated Studies in Diversified Research Contexts

As this study was conducted in an EFL context, more studies are recommended in different instructional environments, especially in L1 and ESL contexts, to verify the effects of peer feedback and back-feedback on writing performance. It is vital to discover the role back-feedback plays in writing practices, as little empirical study about back-feedback has been reported. Though this study sheds general light on the effect of back-feedback, more replicated studies are needed to verify its validity and applicability in diversified instructional environments.

In-depth Exploration into Sub-aspects

This research analyzed the overall quality of feedback and back-feedback only, and thus lacked in-depth investigation into sub-aspects of peer feedback and back-feedback. Though the peer feedback and back-feedback guidance sheets consist of eight aspects of writing (six global, one local, and one summary), this research utilized only the overall scores of all these aspects without examining sub-scores for each aspect. Some replicated studies are needed to further analyze sub-aspects of feedback and back-feedback to examine which aspects facilitate writing achievement most and which ones need more concern.

Small Scale Qualitative Studies

Focusing solely on quantitative data analysis, this research left students' qualitative comments unexplored. Some replicated small scale qualitative studies are recommended to probe into the text of each sub-aspect (six global, one local, and one summary) in the peer feedback and back-feedback guidance sheets to analyze the relationship between each sub-aspect and writing performance. Detailed qualitative analysis may shed further insight into specific instructional designs for future peer interactive studies.

Longitudinal Studies

This study, lasting only 15 weeks and centering exclusively on argumentative essays, may not be long enough to fully reflect upon the effectiveness of peer feedback and back-feedback as reviewing and revision abilities require a sufficient amount of time to internalize (Min, 2006). For more solid evidence, a longitudinal study is highly recommended, which aims at the long-term and overall effectiveness of peer review on writing achievement through exploration into different genres and over time.

Relationship between Peer Review and Self-review

This study indicates that both peer feedback, and back-feedback in particular (a special kind of self-feedback in nature), sharpen students' reflexive ability, or self-review ability. Without a pre-test on students' self-review ability, it is still unclear that to what degree their self-review ability is affected by their feedback activities. Correlational

studies using revised feedback /back-feedback guidance sheets can be conducted to explore how peer review activities influence self-review ability before and after peer feedback training.

Summary

This chapter seeks to discuss and interpret the effects of peer feedback on Chinese English sophomores' writing performance. The results revealed that peer review activities positively and significantly influence students' writing achievement, which supports findings of major current studies on the effectiveness of peer review (Demiraslan Çevik, 2015; Ge, 2011; Li, Liu, & Steckelberg, 2010; Lundstrom & Baker, 2009; Topping, 2009; Zhao, 2014). Further analysis suggests that producing constructive feedback and back-feedback involves and in turn nurtures higher order abilities, which helps to consolidate the reviewers' writing ability (Demiraslan Çevik, 2015; Kim, 2005; Li, Liu, & Steckelberg, 2010; Lundstrom & Baker, 2009).

This study proposes valuable pedagogical recommendations utilizing effective peer training, rubrics and sheets: carefully-designed training processes and well-developed instruments are crucial to the application of peer feedback. It also recommends further exploration including replicated studies in EFL/ESL/L1 contexts, in-depth quantitative studies, small scale qualitative studies, and further exploration of how peer feedback influences self-feedback abilities.

Importantly, this study provides further evidence to apply formative peer feedback in higher education to deal with issues resulting from “larger institutional problems” (Ferris et al., 2011, p. 223), especially, in countries with large populations and academic institutions where the power of formative peer feedback has been mostly ignored (Liu & Carless, 2006). It is very urgent for the decision makers to notice the role formative assessment and feedback plays and to enforce it in the university curriculum with an aim to encourage students to direct and monitor their own learning process and be life-long learners (Gan, 2009; Nicol & Macfarlane-Dick, 2006).

APPENDIX A

IRB APPROVAL LETTER



Institutional Review Board for the Protection of Human Subjects

DATE: November 16, 2017
FROM: Oakland University IRB TO: xin zhang, Ph.D
PROJECT TITLE: An Examination of the Effectiveness of Peer Feedback on Chinese University Students' English Writing Performance
REFERENCE #: 969188-4
SUBMISSION TYPE: Continuing Review/Progress Report
ACTION: APPROVED
APPROVAL DATE: November 16, 2017
EXPIRATION DATE: November 15, 2018
REVIEW TYPE: Expedited Review
REVIEW CATEGORY: Expedited review category # 7
IRB MEETING DATE: November 30, 2017

Thank you for your submission of Continuing Review/Progress Report materials for this project. The Oakland University IRB has APPROVED your submission. This approval is based on an appropriate risk/ benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission package includes the following approved document:

- Continuing Review Application

This research remains active for data analysis only.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the project and assurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the project via a dialogue between I and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure. Do not collect data while the revised application is being reviewed. Data collected during this time cannot be used.

All UNANTICIPATED PROBLEMS involving risks to subjects or others (UPIRSOs) and SERIOUS and UNEXPECTED adverse events must be reported promptly to this committee. Please use the appropriate reporting forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

This project has been determined to be a Minimal Risk project. Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of November 15, 2018.

Please note that all research records must be retained for a minimum of three years after the completion of the project.

Please retain a copy of this correspondence for your record.

If you have any questions, please contact Kate Wydeven M.S. at (248) 370-4306 or kwydeven@oakland.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within Oakland University IRB's records.

APPENDIX B
GUIDANCE SHEETS

Appendix B-1 Feedback Guidance Sheet (for students)

1) What is the writer's thesis statement? Does it clearly express the writer's position? Does the thesis statement contain too much information or any unimportant information? Make a suggestion if the thesis statement is not clear.

2) What are the topic sentences? Do all of these topic sentences logically support the writer's opinion? Please explain why or why not. Make suggestions if the topic sentences are not clear.

3) Identify at least one proper usage of an example and one improper usage (if available) in the paper. Briefly explain each choice and make a suggestion for the improper example.

4) What type of organizational structure (for example, chronological, topical, or logical) does the writer use? Does this order facilitate clarity for the reader? Make a suggestion if the structure is not clear.

5) Does the writer use transitional sentences/words between paragraphs? Identify any instances where the writer lacks a transition and make suggestions.

6) Is there a conclusion? Does it restate the thesis statement in a different way? Does it effectively summarize the writer's position? Why or why not? Make a suggestion if the conclusion is not clear.

7) Are there any grammatical errors or inappropriate word usage? Please only underline the errors and DO NOT correct them directly.

8) What are two strong aspects and two weak aspects of the essay? Please be specific!

Appendix B-2 Back-feedback Guidance Sheet (for students)

1) Which peer suggestions on your thesis statement will you accept and which will you reject? Why?

2) Which peer suggestions on your topic sentences will you accept and which will you reject? Why?

3) Which peer suggestions on your examples will you accept and which will you reject? Why?

4) Which peer suggestions on your organizational structure will you accept and which will you reject? Why?

5) Which peer suggestions on your use of transitions will you accept and which will you reject? Why?

6) Which peer suggestions on your conclusion will you accept and which will you reject? Why?

7) Which peer suggestions on your grammar will you accept and which will you reject? Why?

8) What components (at least 3) do you need to improve in your next draft? Please be specific!

APPENDIX C
SCORING RUBRICS

APPENDIX C-1 Essay Scoring Rubric

Essay Scoring Rubric

Impressive	Score/s
_ Strong organization of essay and paragraphs _ Persuasive reasoning through varied and detailed examples _ Demonstrates style through sophisticated and varied vocabulary, complex grammar and sentence structure, accurate spelling, and effective transitions and punctuation	6
Clearly competent	Score/s
_ Clear organization of essay and paragraphs _ Relevant, detailed examples _ Correct use of most vocabulary, grammar, sentence structure, transitions, spelling, and punctuation; minor errors do not interfere with communication	5
Satisfactory (sometimes only marginally)	Score/s
_ Organized essay and paragraphs _ Developed with adequate examples, but lacking detail _ Correct use of most vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation; occasional errors sometimes interfere with communication	4
Unsatisfactory	Score/s
_ Some evidence of organization of essay and/or paragraphs _ Little development _ Frequent errors in vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation sometimes interfere with communication	3
Weak	Score/s
_ Slight evidence of organization of essay and/or paragraphs, but ideas confused and/or disconnected _ Very little development, but simplistic _ Frequent and varied errors in vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation interfere with communication	2
Severely limited	Score/s
_ No evidence of organization _ No development _ Limited to basic words, phrases, and sentences often with errors _ May be off topic or merely a copy of the Essay Test Topic	1

Appendix C-2 Feedback Scoring Rubric (for raters)

Feedback Scoring Rubric

Student:	Rater:
Thesis statement	Score/s
0 points: No comments or unrelated comments on thesis statement 1 point: Comments on thesis statement given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on thesis statement and good suggestions for improvement	
Topic sentences	Score/s
0 points: No comments or unrelated comments on topic sentences 1 point: Comments on topic sentences given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on topic sentences and good suggestions for improvement	
Usage of examples	Score/s
0 points: No comments or unrelated comments on usage of examples 1 point: Comments on usage of examples given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on usage of examples and good suggestions for improvement	
Organizational structure	Score/s
0 points: No comments or unrelated comments on organizational structure 1 point: Comments on organizational structure given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on organizational structure and good suggestions for improvement	
Transitions	Score/s
0 points: No comments or unrelated comments on transitions 1 point: Comments on transitions given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on transitions and good suggestions for improvement	
Conclusion	Score/s
0 points: No comments or unrelated comments on conclusion 1 point: Comments on conclusion given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on conclusion and good suggestions for improvement	

Grammar	Score/s
0 points: No comments or unrelated comments on grammar 1 point: Comments on grammar given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on grammar and good suggestions for improvement	
Strength and weakness	Score/s
0 points: No comments or unrelated comments on strength and weakness 1 point: Comments on strength and weakness given and suggestions made for improvement, but feedback not clear 2 points: Clear comments on strength and weakness and good suggestions for improvement	
Word requirement	Score/s
0 points: Less than 150 words 1 point: 150 words and more	
Total	/17

Appendix C-3 Back-feedback Scoring Rubric (for raters)

Back-feedback Scoring Rubric

Student:	Rater:
Thesis statement	Score/s
0 points: No explanation of acceptance or rejection of suggestions on thesis statement 1 point: Explanation of acceptance or rejection of suggestions on thesis statement given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on thesis statement	
Topic sentences	Score/s
0 points: No explanation of acceptance or rejection of suggestions on topic sentences 1 point: Explanation of acceptance or rejection of suggestions on topic sentences given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on topic sentences	
Usage of examples	Score/s
0 points: No explanation of acceptance or rejection of suggestions on usage of examples 1 point: Explanation of acceptance or rejection of suggestions on usage of examples given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on usage of examples	

Organizational structures	Score/s
0 points: No explanation of acceptance or rejection of suggestions on organizational structures 1 point: Explanation of acceptance or rejection of suggestions on organizational structures given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on organizational structures	
Transitions	Score/s
0 points: No explanation of acceptance or rejection of suggestions on transitions 1 point: Explanation of acceptance or rejection of suggestions on transitions given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on transitions	
Conclusion	Score/s
0 points: No explanation of acceptance or rejection of suggestions on conclusion 1 point: Explanation of acceptance or rejection of suggestions on conclusion given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on conclusion	
Grammar	Score/s
0 points: No explanation of acceptance or rejection of suggestions on grammar 1 point: Explanation of acceptance or rejection of suggestions on grammar given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on grammar	
Components needed to improve in the next draft	Score/s
0 points: No explanation of acceptance or rejection of suggestions on components needed to improve in the next draft 1 point: Explanation of acceptance or rejection of suggestions on components needed to improve in the next draft given, but not clear 2 points: Clear explanation of acceptance or rejection of suggestions on components needed to improve in the next draft	
Word requirement	Score/s
0 point: Less than 150 words 1 point: 150 words and more	
Total	/17

APPENDIX D

PRE-AND POST-TEST PROMPTS

Pre-test Prompt**[45 MIN]**

Should we revive traditional Chinese characters or continue using simplified characters? This has been an intensely discussed question for years. The following are the supporters' and opponents' opinions. Read carefully the opinions from both sides and write your response in no less than **200** words, in which you should:

1. summarize the opinions from both sides, and then
2. give your view on the issue.

Marks will be awarded for content relevance, content sufficiency, organization and language quality. Failure to follow the above instructions may result in a loss of marks.

Yes	No
Traditional characters, which date back to more than 2000 years ago, have a more beautiful appearance and a more reasonable structure. As indicated by the 親 and 愛 examples, traditional characters make more sense, convey traditional values and can therefore represent traditional culture. For two millennia, Chinese historical records and classic works were written in traditional characters. To be able to read them and inherit traditional culture, we need to bring traditional characters back. Politically, it is also necessary to restore traditional Chinese characters. Currently, traditional characters are still in used in Hong Kong, Taiwan and many Chinese communities around the world. Restoring them can contribute to cross-Straits exchanges and national reunification and unite Chinese people around the world.	In today's world, efficiency matters most. Traditional characters, which usually have more strokes than simplified ones, are more difficult to learn. By contrast, simplified characters are much easier to learn and use. Over the past 50 years, lots of classic texts have been turned into simplified-character versions, which means simplified characters can also promote and preserve traditional culture. Constant simplification has been a trend in the evolution of Chinese characters. From the oracle bones script of 3000 years ago to traditional characters, the Chinese writing system has always been slimming down for better communication.

Post-test Prompt**[45 MIN]**

Could online courses be an effective alternative for the traditional in-class course in university education? With the continued growth of online teaching systems and integration of massive open online courses (MOOCS) into higher education, college study will never be the same for both professors and students. The following are opinions from both sides. Read the excerpts carefully and write your response in no less than **200** words, in which you should:

1. summarize the opinions from both sides, and then
2. give your view on the issue.

Marks will be awarded for content relevance, content sufficiency, organization and language quality. Failure to follow the above instructions may result in a loss of marks.

Yes	No
Universities: The University of Washington (UW) in the US concedes online courses may be more effective for self-directed learners. Student1, Jasmine Barta of Arizona State University: Online classes are the secret to a happier, fuller and less stressful college experience. Student 2: Chang Hanyi of Boston University: The coursework is actually as demanding as my other regular language classes. Despite the hard work, I still enjoy cyber interaction with my professor and classmates. Student 3: Yang Yang of Peking University: With MOOCS, we are no longer confined to a classroom at a certain time slot.	Universities: The University of Washington (UW) in the US concedes students who are not organized and in possession of good time-management skills may struggle. UW also mentions online courses may not be able to accurately replicate the vibrant sense of intellectual community that has been at the heart of higher education for ages. Student1: Jasmine Barta of Arizona State University: Some students complain about the lack of social interaction and the ease with which they can forget to meet a deadline. Student 2: Chang Hanyi of Boston University: My professor even invited me to face-to-face meetings four times to address article structure in my writing assignments. So, taking online courses doesn't mean zero physical interaction with your instructors. Student 3, Yang Yang of Peking University: Although MOOCs cannot compete with traditional ways of learning in terms of teacher-student interaction. I used to ask questions immediately after class. But with MOOCs, we only have office hours for question and answer sessions. Most of the time, I won't bother to go.

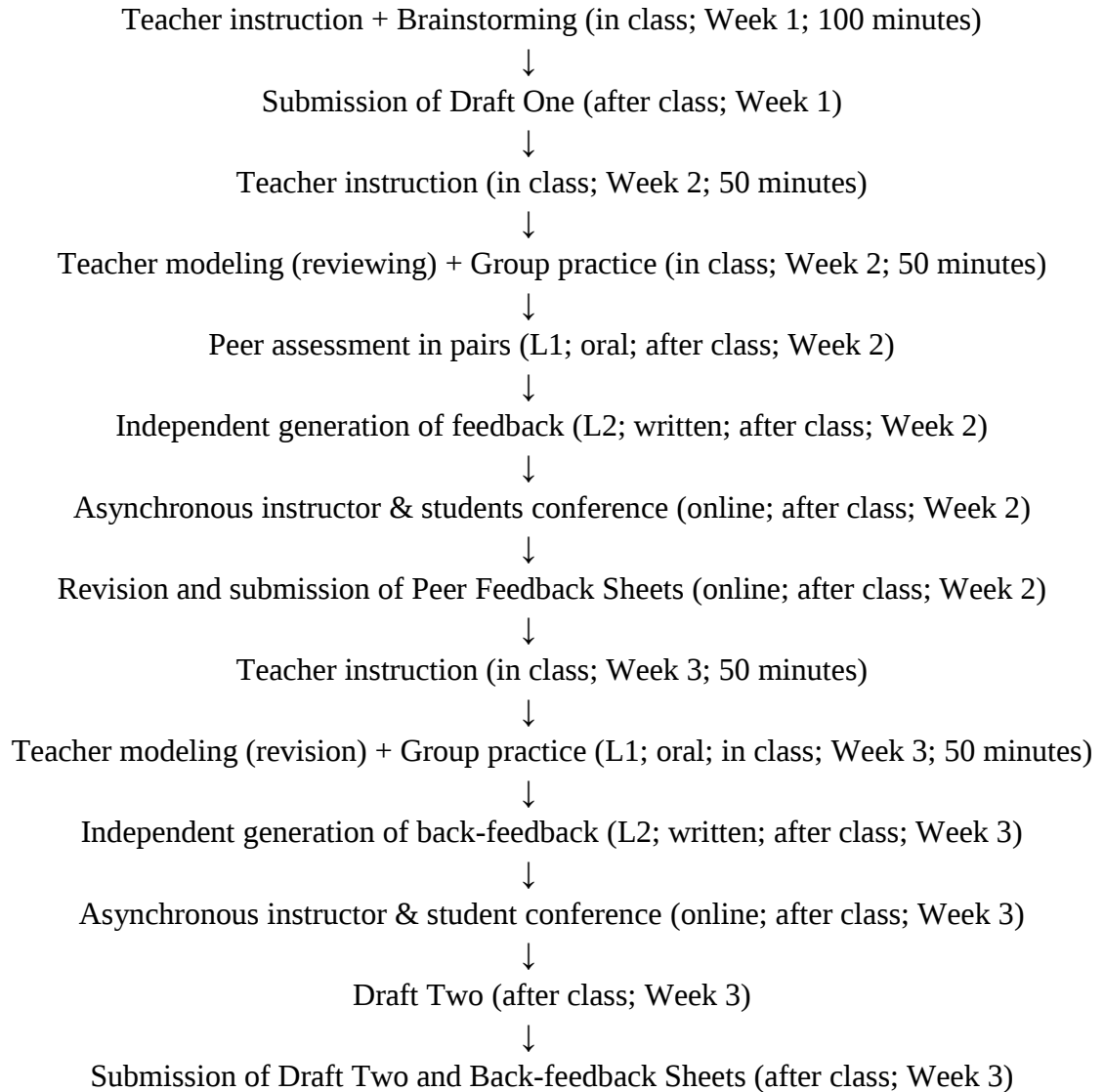
APPENDIX E
THE RESEARCH SCHEDULE

The research schedule (14 weeks, 100 minutes per week)

Week	Task	Content and activities	Note
1	Course introduction	Syllabus discussion; consent form collected	Same for both groups
2	Pre-test Teacher instruction	1.Pretest (45 minutes; both control and experimental groups) 2.Initial persuasion (experimental group)	1.Pre-test compositions collected 2. Teacher instruction for control group while experimental group receives initial persuasion training
3	1.Teacher instruction 2. Writing 1	1.Writing skills 2.Pre-writing activities (in class) 3.First draft (out of class)	Same for both groups Digital first drafts collected
4	1.Teacher instruction 2.Training on reviewing process using feedback guidance sheet	1.Writing skills 2.Teacher's modeling of reviewing process 3. Teacher's explanation of feedback scoring rubric 4.Group reviewing (1-4 in class) 5. Reviewing and generating feedback individually 6. Online teacher-student conference 7. Submission of feedback sheet (5-7 out of class)	1.Control group skips steps 2-4, but do in-class writing and group discussion 2. Control group reviews what has been discussed in class and do some exercises while experimental group does steps 5-7
5	1.Teacher instruction 2. Training on revision process using Back-feedback guidance sheet	1.Writing skills 2.Teacher's modeling of revision process 3. Teacher's explanation of Back-feedback scoring rubric 4. Group revision (1-4 in class) 5. Revising and generating Back-feedback individually 6. Online teacher-student conference 7. Submission of second drafts and Back-feedback sheet (5-7 out of class)	1. Control group skips steps 2-4, but do in-class writing and discussion 2. Control group reviews teacher feedback and writes second drafts while experimental group does steps 5-7 3. Digital second drafts collected
6	1.Teacher instruction 2. Writing 2	Steps 1-3 in week 3	Same steps in week 3
7	Steps 1-2 in week 4	Steps 1-7 in week 4	Same steps in week 4
8	Steps 1-2 in week 5	Steps 1-7 in week 5	Same steps in week 5
9	1.Teacher instruction 2. Writing 3	Steps 1-3 in week 3	Same step in week 3
10	1. Steps 1-2 in week 5	Steps 1-7 in week 4	Same steps in week 4
11	1. Steps 1-2 in week 6	Steps 1-7 in week 5	Same steps in week 5
12	1.Teacher instruction 2. Writing 4	Steps 1-3 in week 3	Same step in week 3
13	1. Steps 1-2 in week 5	Steps 1-7 in week 4, excluding step 6	Same steps in week 4
14	1. Steps 1-2 in week 6	Steps 1-7 in week 5, excluding step 6	Same steps in week 5
15	1. Post-test 2. Course evaluation	1. Posttest (45 minutes) 2. Course evaluation	Post-test compositions collected

APPENDIX F

ONE WRITING CYCLE FOR EXPERIMENTAL GROUP (3 WEEKS)



APPENDIX G

SAMPLES

Appendix G-1 Sample Training Essay

Should Lady-boys Be Allowed on TV

1. Nowadays the so-called lady-boys are talked by many people. In China this phenomenon became such a hot topic due to a girl-like boy appeared on the entertainment show in Hunan TV. Should such people be accepted in our society? The answer is certain. But should such people be allowed on the screen in front of billions of people? Well, I don not think so. To me, it has bad influences.

2. Lady-boys is a kind of group like gays. We should accept their ways of life. Because they deserve it just like other humans. Firstly, as we all know, lady-boys appeared a long time ago. Secondly, the society allows people like this to show themselves to others through clubs or other ways.

3. But they should not be allowed on TV. There are mainly two reasons.

4. First, if lady-boys is allowed on TV, they may have great influence to the young generations. TV is very easy attached to young children and children may learn from what they saw on the TV show. They may think that lady-boys is popular in our society and learn to act as them, which is sick to the normal development. Secondly, lady-boys bent people's concept to aesthetics. Lady-boys is a group that doesn't naturally normal. Compared to such entertainment like lady-boys, people like the traditional means of entertainment. Because the traditional entertainment is what they think normal.

5. In conclusion, lady-boys should not be allowed on TV. But views on lady-boys should be judged differently. People should change their old thoughts about lady-boys since we are now in a world that accepts diversification. (269 words)

Appendix G-2 Sample Feedback

1) What is the writer's thesis statement? Does it clearly express the writer's position? Does the thesis statement contain too much information or any unimportant information? Make a suggestion if the thesis statement is not clear.

There is not a clear thesis statement at the end of the first paragraph. If you can provide a specific thesis statement, you can better guide your audience to what you are going to talk about later. For instance, you could consider saying that lady-boys should not be allowed on TV because their appearance on TV may lead to some bad influences.

2) What are the topic sentences? Do all of these topic sentences logically support the writer's opinion? Please explain why or why not. Make suggestions if the topic sentences are not clear.

There are no topic sentences in the second and the fourth paragraphs. As you mentioned, "it has bad influences" at the end of your first paragraph. If you can split your fourth paragraph into two separate paragraphs and mention one of the bad influences in your topic sentence of each paragraph, your organization will be clearer.

3) Identify at least one proper usage of an example and one improper usage (if available) in the paper. Briefly explain each choice and make a suggestion for the improper example.

Your opinion is "lady-boys should not be allowed on TV". However, in your second paragraph, you talk about "We should accept their ways of life." this seems to be unrelated to your opinion. Another example might be better.

4) What type of organizational structure (for example, chronological, topical, or logical) does the writer use? Does this order facilitate clarity for the reader? Make a suggestion if the structure is not clear.

There seems to be no clear organizational structure of the essay. If you can indicate the order of the paragraphs according to the importance of the influences or the logical order in your topic sentences or thesis statement, it would make more sense.

5) Does the writer use transitional sentences/words between paragraphs? Identify any instances where the writer lacks a transition and make suggestions.

The third paragraph serves as a transition, but it might be too short to be a separate paragraph. These two sentences would be better if they are put at the end of the second paragraph. In other paragraphs such as the fourth one, you make good use of transitions, like “first” and “second.”

6) Is there a conclusion? Does it restate the thesis statement in a different way? Does it effectively summarize the writer’s position? Why or why not? Make a suggestion if the conclusion is not clear.

You restate your position in the last paragraph, but your next two sentences don’t support your conclusion. You should summarize your opinions here.

7) Are there any grammatical errors or inappropriate word usage? Please only underline the errors and DO NOT correct them directly.

There are some minor mistakes. For instance, there is one mistake in the first sentence of the second and fourth paragraphs.

8) What are two strong aspects and two weak aspects of the essay? Please be specific!

In your first paragraph, it is very good that you state your position clearly. It is also good that you use transitions such as “first” and “second” to guide your audience. One of your weaknesses might be that certain examples selected do not support your opinion, such as the second paragraph. Another potential weakness is that you don’t provide clear topic sentences.

Appendix G-3 Sample Back-feedback

1) Which peer suggestions on your thesis statement will you accept and which will you reject? Why?

I will accept the suggestion regarding a missing thesis statement because I only described the phenomenon in my first paragraph without providing a clear thesis statement. As suggested by the reviewer, if I mention that there will be bad influences in my thesis statement, I can better guide my audience to what I am going to talk about later.

2) Which peer suggestions on your topic sentences will you accept and which will you reject? Why?

I will accept the suggestion regarding missing topic sentences. If I can mention one of these bad influences specifically at the beginning of each paragraph, it will help me to reflect on what my focus is and my readers can better understand me.

3) Which peer suggestions on your examples will you accept and which will you reject? Why?

I will accept the suggestion regarding my inappropriate example selection in the second paragraph because it doesn't support my position.

4) Which peer suggestions on your organizational structure will you accept and which will you reject? Why?

I will accept the suggestion regarding unclear organizational structure. I think the suggestion of indicating the order using topic sentences is a good idea.

5) Which peer suggestions on your use of transitions will you accept and which will you reject? Why?

I will not accept the suggestion of deleting a short paragraph as a transition. I think it is emphatic to use short sentences or paragraph as a transition between paragraphs.

6) Which peer suggestions on your conclusion will you accept and which will you reject? Why?

I will accept the suggestion on my conclusion because it does not support my opinion. I think I am off topic in the ending and I need to summarize my ideas here.

7) Which peer suggestions on your grammar will you accept and which will you reject? Why?

I will accept the suggestion on my grammar because these suggestions are correct. The subjects do not match the predicates in the sentences that were pointed out.

8) What components (at least 3) do you need to improve in your next draft? Please be specific!

I will split my fourth paragraph into two separate sections and mention one of the bad influences in my topic sentence of each paragraph. I will provide a specific thesis statement at the end of the first paragraph. I will get rid of the second paragraph because it does not support my opinion, and I will rewrite my ending because it doesn't summarize my opinions.

Appendix G-4 Sample Language Structures

Thesis statement
Your thesis statement is concise/ can be more concise If you can/could provide a specific thesis statement, you can/could For instance, you could consider saying...
Topic sentences
Your topic sentence is concise/ can be more concise For instance, you could consider saying... It could/would make more sense, if you can ...
Usage of examples
Your opinion is ... However, you talk about... this seems to be unrelated to your opinion. These/this example(s) strongly support(s)/ does not support your idea. I am not sure... Can you clarify this for me? What do you mean by saying/ using?
Organizational structure
There seems to be no clear organizational structure of the essay. If you can indicate the order..., it would make more sense. It would be better, if you put...at the end/beginning of (paragraph)
Transitions
You make good use of transitions in You may need transitions such as ... in paragraph(s)...
Conclusion
You restate your position in the last paragraph, but... You may need to summarize your opinion Your conclusion is clear and strong... If you can/could..., your conclusion can be stronger.
Grammar
There are some grammatical mistakes, but they are minor. For instance,
Strength and weakness
It is very clear /good/impressive that you ... It is also clear/good/impressive that you ... One of your weaknesses might be... The other might be
Word requirement
Your essay satisfies (does not satisfy) the word requirement.
Accepting or rejecting suggestions
I will accept the suggestion regarding, because I may not accept the suggestion regarding, because

APPENDIX H

FIGURES

Normal P-P Plot of Regression Standardized Residual

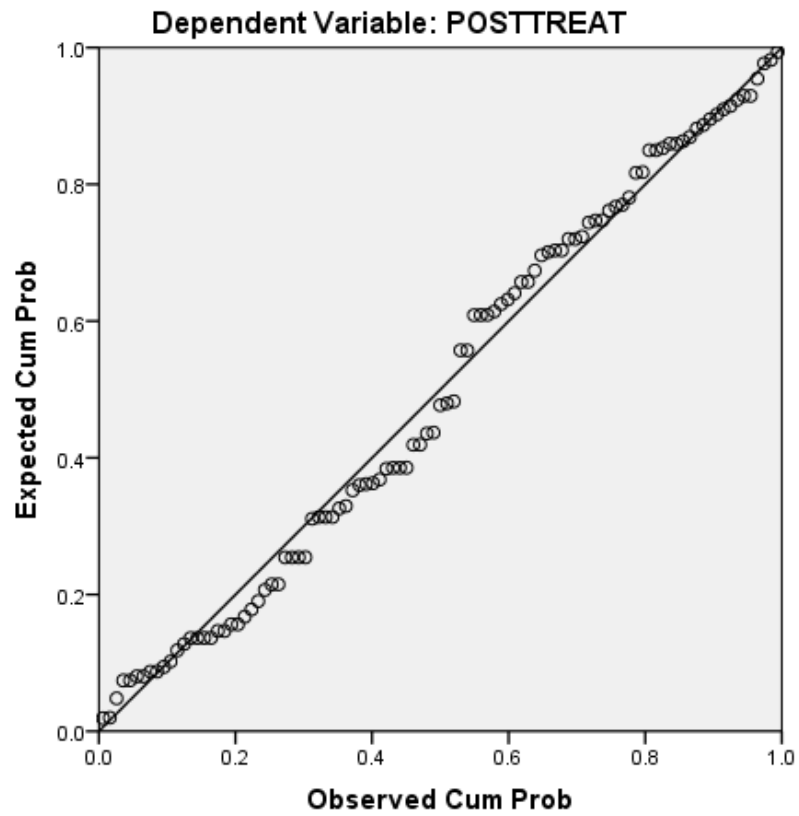


Figure H-1: Normal P-P Plot of Regression Standardized Residual

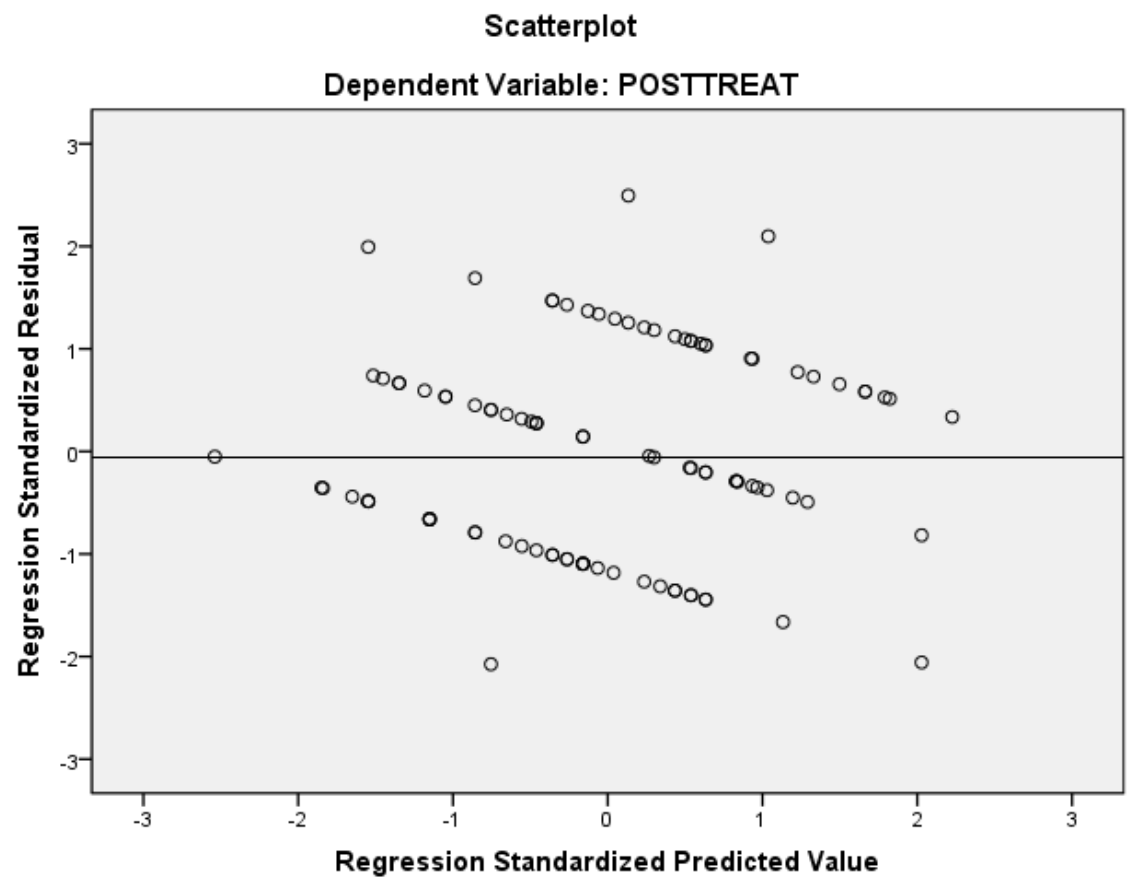


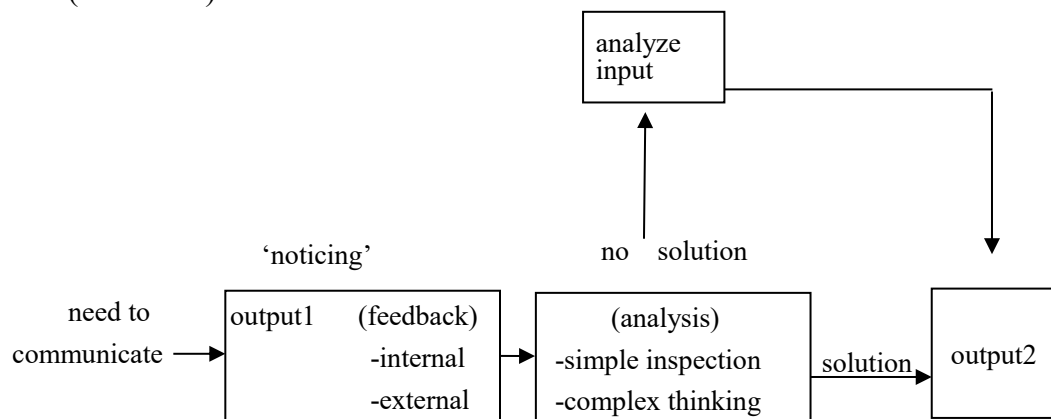
Figure H-2: Scatterplot

APPENDIX I
COPYRIGHT PERMISSION

This license below will confirm our recent Email conversation and telephone conversation regarding the reprinting of a figure in my dissertation, which I am currently completing at Oakland University in Rochester, Michigan. My dissertation is entitled "AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON CHINESE UNIVERSITY STUDENTS' ENGLISH WRITING PERFORMANCE." I would like your permission to reprint in my dissertation excerpts from the following:

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391. <https://doi-org.huaryu.kl.oakland.edu/10.1093/applin/16.3.371>

The figure to be reproduced is: Output and second language learning (see below):



OXFORD UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS

Jan 03, 2018

This Agreement between xin zhang ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number 4257830888013

License date	Dec 28, 2017
Licensed content publisher	Oxford University Press
Licensed content publication	Applied Linguistics
Licensed content title	Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning
Licensed content author	SWAIN, MERRILL; LAPKIN, SHARON
Licensed content date	Sep 1, 1995
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON CHINESE UNIVERSITY STUDENTS' ENGLISH WRITING PERFORMANCE
Publisher of your work	n/a
Expected publication date	Apr 2018
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Requestor Location	xin zhang 2656 lantern lane, 206 AUBURN HILLS, MI 48326 United States Attn: xin zhang
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing Address	xin zhang 2656 lantern lane, 206 AUBURN HILLS, MI 48326, United States
This license below will confirm our recent Email conversation and telephone conversation regarding the reprinting of an excerpt in my dissertation, which I am currently completing at Oakland University in Rochester, Michigan. My dissertation is entitled "<u>AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON CHINESE</u>	

UNIVERSITY STUDENTS' ENGLISH WRITING PERFORMANCE." I would like your permission to reprint in my dissertation excerpts from the following:

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179-200. doi:10.1016/j.jslw.2006.09.004

The excerpt to be reproduced is the essay scoring guide (see below)

6. Impressive
 - Strong organization of essay and paragraphs
 - Persuasive reasoning through varied and detailed examples
 - Demonstrates style through sophisticated and varied vocabulary, complex grammar and sentence structure, accurate spelling, and effective transitions and punctuation
5. Clearly competent
 - Clear organization of essay and paragraphs
 - Relevant, detailed examples
 - Correct use of most vocabulary, grammar, sentence structure, transitions, spelling, and punctuation; minor errors do not interfere with communication
4. Satisfactory (sometimes only marginally)
 - Organized essay and paragraphs
 - Developed with adequate examples, but lacking detail
 - Correct use of most vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation; occasional errors sometimes interfere with communication
3. Unsatisfactory
 - Some evidence of organization of essay and/or paragraphs
 - Little development
 - Frequent errors in vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation sometimes interfere with communication
2. Weak
 - Slight evidence of organization of essay and/or paragraphs, but ideas confused and/or disconnected
 - Very little development, but simplistic
 - Frequent and varied errors in vocabulary, grammar, sentence structure, transitions, spelling, and/or punctuation interfere with communication
1. Severely limited
 - No evidence of organization
 - No development
 - Limited to basic words, phrases, and sentences often with errors
 - May be off topic or merely a copy of the Essay Test Topic

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Jan 03, 2018

This Agreement between xin zhang ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4258831390543
License date	Dec 30, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Second Language Writing
Licensed Content Title	A comparative study of peer and teacher feedback in a Chinese EFL writing class
Licensed Content Author	Miao Yang, Richard Badger, Zhen Yu
Licensed Content Date	Dec 1, 2006
Licensed Content Volume	15
Licensed Content Issue	3
Licensed Content Pages	22
Start Page	179
End Page	200
Type of Use	reuse in a thesis/dissertation
Portion	excerpt
Number of excerpts	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

Order reference number	1
Title of your thesis/dissertation	AN EXAMINATION OF THE EFFECTIVENESS OF PEER FEEDBACK ON CHINESE UNIVERSIRY STUDENTS' ENGLISH WRITING PERFORMANCE
Expected completion date	Apr 2018
Estimated size (number of 180 pages)	
Requestor Location	xin zhang 2656 lantern lane, 206 AUBURN HILLS, MI 48326 United States Attn: xin zhang
Publisher Tax ID	98-0397604

REFERENCES

- Badger, R., & White, G. (2000). A process genre approach to teaching writing. *ELT Journal*, 54(2), 153-160. doi:10.1093/elt/54.2.153
- Barrett, P. (2001). Assessing the reliability of rating data. Retrieved from <https://pdfs.semanticscholar.org/41e3/46312970950a2c6690cb16a73b938d1c6cfc.pdf>
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241. doi:10.1016/S1060-3743(99)80115-5
- Biggs, J. & Tang, C. (2007). *Teaching for quality learning at university: what the student does*. (3rd ed.). Philadelphia, Pa.: Society for Research into Higher Education. Open University Press.
- Bruffee, K. A. (1984). Collaborative learning and the "conversation of mankind". *College English*, 46(7), 635-652. doi: 10.2307/376924
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245-281. doi:10.2307/1170684
- Carson, J. G., & Nelson, G. L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1-19. doi:10.1016/S1060-3743(96)90012-0
- Chang, C. Y. (2016). EFL reviewers' emoticon use in asynchronous computer-mediated peer response. *Computers and Composition*, 40, 1-18. doi:10.1016/j.compcom.2016.03.008
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409-426. doi:10.1016/j.compedu.2005.02.004
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. doi:http://dx.doi.org.huaryu.kl.oakland.edu/10.1037/0033-2909.112.1.155

- Connor, U. & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing*, 3, 257-276.
[https://doi-org.huaryu.kl.oakland.edu/10.1016/1060-3743\(94\)90019-1](https://doi-org.huaryu.kl.oakland.edu/10.1016/1060-3743(94)90019-1)
- Crinon, J., & Marin, B. (2010). The role of peer feedback in learning to write explanatory texts: Why the tutors learn the most. *Language Awareness*, 19(2), 111-128.
 doi:10.1080/09658411003746604
- Davydov, V. V., & Kerr, S. T. (1995). The influence of L. S. Vygotsky on education theory, research, and practice. *Educational Researcher*, 24(3), 12-21.
 doi:10.2307/1176020
- De Guerrero, M. C. M., & Villamil, O. S., (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal*, 84(1), 51-68.
 doi:10.1111/0026-7902.00052
- Demiraslan Çevik, Y. (2015). Assessor or assessee? Investigating the differential effects of online peer assessment roles in the development of students' problem-solving skills. *Computers in Human Behavior*, 52, 250-258. doi:10.1016/j.chb.2015.05.056
- Diab, N. M. (2011). Assessing the relationship between different types of student feedback and the quality of revised writing. *Assessing Writing*, 16(4), 274-292.
 doi:10.1016/j.asw.2011.08.001
- Driscoll, D. L. (2011, December). Connected, disconnected, or uncertain: Student attitudes about future writing contexts and perceptions of transfer from first year writing to the disciplines. *Across the Disciplines*, 8(2). Retrieved June 22, 2013, from <http://wac.colostate.edu/atd/articles/driscoll2011/index.cfm>
- Falchikov, N. (2005). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. New York: RoutledgeFalmer
- Farris, C. (1987). Current Composition: Beyond Process vs. Product. *The English Journal*, 76(6), 28-34. doi:10.2307/818051
- Ferris, D. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, 31, 315-339. doi: 10.2307/3588049

- Ferris, D., Brown, J., Liu, H., & Stine, M. E. A. (2011). Responding to L2 students in college writing classes: Teacher perspectives. *TESOL Quarterly*, 45(2), 207-234. doi:10.5054/tq.2011.247706
- Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, diagnosis, and the strategies of revision. *College Composition and Communication*, 37(1), 16-55. doi: 10.2307/357381
- Foley, S. (2013). Student views of peer assessment at the international school of lausanne. *Journal of Research in International Education*, 12(3), 201-213. doi:10.1177/1475240913509766
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2015). *How to design and evaluate research in education*. New York, NY: McGraw-Hill Education
- Gan, Z. (2009). "Asian learners" re-examined: An empirical study of language learning attitudes, strategies and motivation among mainland Chinese and Hong Kong students. *Journal of Multilingual and Multicultural Development*, 30(1), 41-58. doi:10.1080/01434630802307890
- Ge, Z. (2011). Exploring e-learners' perceptions of net-based peer-reviewed English writing. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 75-91. doi:10.1007/s11412-010-9103-7
- Gielen, M., & De Wever, B. (2015). Structuring the peer assessment process: A multilevel approach for the impact on product improvement and peer feedback quality. *Journal of Computer Assisted Learning*, 31(5), 435-449. doi:10.1111/jcal.12096
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20(4), 304-315. doi:10.1016/j.learninstruc.2009.08.007
- Graham, S., & Harris, K. R. (2013). Common core state standards, writing, and students with LD: Recommendations. *Learning Disabilities Research & Practice*, 28(1), 28-37. doi:10.1111/ldrp.12004
- Graham, S., & Harris, K. R. (2016). A path to better writing. *The Reading Teacher*, 69(4), 359-365. doi:10.1002/trtr.1432

- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445-476.
doi:10.1037/0022-0663.99.3.445
- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443-461.
doi:10.1016/j.compcom.2007.03.002
- Hansen, J. G., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal*, 59(1), 31-38. doi:10.1093/elt/cci004
- Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, 3(2), 141-163. doi:10.1016/1060-3743(94)90012-4
- Ho, M. C. (2015). The effects of face-to-face and computer-mediated peer review on EFL writers' comments and revisions. *Australasian Journal of Educational Technology*, 31(1), 1-15. <https://doi.org/10.14742/ajet.495>
- Ho, R. (2013). *Handbook of univariate and multivariate data analysis with IBM SPSS* (2nd ed.). Hoboken: CRC Press.
- Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development*, 27(1), 55-67.
doi: 10.1080/07294360701658765
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), 321-342. doi:10.1191/1362168805lr169oa
- Hu, G., & Lam, S. T. E. (2010). Issues of cultural appropriateness and pedagogical efficacy: Exploring peer review in a second language writing class. *Instructional Science*, 38(4), 371-394. doi:10.1007/s11251-008-9086-1
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236-248. doi:http://dx.doi.org.huaryu.kl.oakland.edu/10.1037/0012-1649.27.2.236
- Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System*, 31(2), 217-230. doi:10.1016/S0346-251X(03)00021-6

- Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and issues*. Cambridge: Cambridge University Press.
- Jacobs, G. M., Curtis, A., Braine, G., & Huang, S. (1998). Feedback on student writing: Taking the middle path. *Journal of Second Language Writing*, 7(3), 307-317. doi:10.1016/S1060-3743(98)90019-4
- Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23(2), 12. <https://doi.org/10.18806/tesl.v23i2.53>
- Keh, C. L. (1990). Feedback in the writing process: a model and methods for implementation. *ELT Journal*, 44, 294-394. <https://doi.org/10.1093/elt/44.4.294>
- Kellogg, R. T., Whiteford, A. P., & Quinlan, T. (2010). Does automated feedback help students learn to write? *Journal of Educational Computing Research*, 42(2), 173-196. doi:10.2190/EC.42.2.c
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., . . . Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, 68(3), 350-386. doi:10.2307/1170601
- Kim, M. (2005). *The effects of the assessor and assessee's roles on preservice teachers' metacognitive awareness, performance, and attitude in a technology -related design task* (Order No. 3198230). Available from ProQuest Dissertations & Theses Global. (304997538). Retrieved from <https://search-proquest-com.huaryu.kl.oakland.edu/docview/304997538?accountid=12924>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. doi:10.1016/j.jcm.2016.02.012.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Lam, R. (2010). A peer review training workshop: Coaching students to give and evaluate peer feedback. *TESL Canada Journal*, 27(2). <https://doi.org/10.18806/tesl.v27i2.1052>

- Lee, G., & Schallert, D. L. (2008a). Meeting in the margins: Effects of the teacher–student relationship on revision processes of EFL college students taking a composition course. *Journal of Second Language Writing*, 17(3), 165-182. doi:10.1016/j.jslw.2007.11.002
- Lee, G., & Schallert, D. L. (2008b). Constructing trust between teacher and students through feedback and revision cycles in an EFL writing classroom. *Written Communication*, 25(4), 506-537. doi:10.1177/0741088308322301
- Lee, I. (2014). Revisiting teacher feedback in EFL writing from sociocultural perspectives. *TESOL Quarterly*, 48(1), 201-213. doi:10.1002/tesq.153
- Leki, I. (1990). Potential problems with peer responding in ESL writing classes. *CATESOL Journal*, 3, 5-19. Retrieved from <https://oakland.illiad.oclc.org/illiad/illiad.dll?Action=10&Form=75&Value=293903>
- Leys, C., Klein, O., Dominicy, Y., & Ley, C. (2018). Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, 74, 150-156. <https://doi.org/10.1016/j.jesp.2017.09.011>
- Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology*, 41(3), 525-536. doi:10.1111/j.1467-8535.2009.00968.x
- Li, L., Liu, X., & Zhou, Y. (2012). Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *British Journal of Educational Technology*, 43(3), 376-384. doi:10.1111/j.1467-8535.2011.01180.x
- Liang, M. (2010). Using synchronous online peer response groups in EFL writing: Revision-related discourse. *Language, Learning & Technology*, 14(1), 45-64. Retrieved from <http://llt.msu.edu/vol14num1/liang.pdf>
- Liou, H., & Peng, Z. (2009). Training effects on computer-mediated peer review. *System*, 37(3), 514-525. doi:10.1016/j.system.2009.01.005
- Liu, J., & Hansen Edwards, J. (2002). Ebooks Corporation. *Peer response in second language writing classrooms*. Ann Arbor: University of Michigan Press.

- Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3), 193-227. doi:10.1016/S1475-1585(03)00025-0
- Liu, N.F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
https://doi-org.huaryu.kl.oakland.edu/10.1080/13562510600680582
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45(4), 605-651.
doi:10.1111/j.1467-1770.1995.tb00456.x
- Long, M. H. (1983b). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126-141.
http://doi-org.huaryu.kl.oakland.edu/10.1093/applin/4.2.126
- Lu, J., & Law, N. (2012). Online peer assessment: Effects of cognitive and affective feedback. *Instructional Science*, 40(2), 257-275. doi:10.1007/s11251-011-9177-2
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30-43. doi:10.1016/j.jslw.2008.06.002
- Ma, J. (2010). Chinese EFL learners' decision-making while evaluating peers' texts. *International Journal of English Studies*, 10(2), 99-120.
doi:10.6018/ijes.10.2.119221
- Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly*, 28(4), 745-769. doi: 10.2307/3587558
- Min, H. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293-308. doi:10.1016/j.system.2004.11.003
- Min, H. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118-141.
doi:10.1016/j.jslw.2006.01.003
- Min, H. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes*, 27(3), 285-305.
doi:10.1016/j.esp.2008.02.002

- Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, 16(2), 82-99. doi:10.1016/j.jslw.2007.04.002
- Moore, L. K. (1986). Teaching students how to evaluate writing. *TESOL Newsletter* 20, 23-24. Retrieved from <https://oakland.illiad.oclc.org/illiad/illiad.dll?Action=10&Form=75&Value=293905>
- Mulder, R. A., Pearce, J. M., & Baik, C. (2014). Peer review in higher education: Student perceptions before and after participation. *Active Learning in Higher Education*, 15(2), 157-171. doi:10.1177/1469787414527391
- Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi:10.1080/03075070600572090
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122. doi:10.1080/02602938.2013.795518
- Nunan, D. (2001). *Second language teaching and learning*. Beijing: Foreign Language Teaching and Research Press.
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36 (2), 89-101. https://doi-org.huaryu.kl.oakland.edu/10.1207/S15326985EP3602_4
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265-289. doi:10.1016/S1060-3743(99)80117-9
- Posten, H.O. (1984). Robustness of the two-sample t-test. In D. Rasch and M.L.Tiku (Eds.), *Robustness of Statistical Methods and Nonparametric Statistics* (pp. 92-99). Dordrecht, Germany: Reidel. doi: 10.1007/978-94-009-6528-7_23
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training : Quality, styles, and preferences. *Advances in Health Sciences Education*, 11(3), 289-303. doi:10.1007/s10459-005-3250-z

- Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10(4), 277-303. doi:10.1016/S1060-3743(01)00046-7
- Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research*, 17(1), 67-89. doi:10.1177/1362168812459151
- Ramsey, P. H. (1980). Exact type 1 error rates for robustness of student's t test with unequal variances. *Journal of Educational Statistics*, 5(4), 337-349. doi:10.3102/10769986005004337
- Ren, H. W., & Hu, G. W. (2012). Peer review and Chinese EFL/ESL student writers. *English Australia Journal*, 27(2), 3-16. Retrieved from <https://repository.nie.edu.sg/bitstream/10497/16145/1/EAJ-27-2-3.pdf>
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59(1), 23-30. doi:10.1093/elt/cci003
- Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, 73-82. doi:10.1016/j.linged.2014.12.001
- Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8(1), 31-54. doi:10.1191/1362168804lr133oa
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158. <https://doi-org.huaryu.kl.oakland.edu/10.1093/applin/11.2.129>
- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust?: Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology*, 6(4), 147-151. doi:10.1027/1614-2241/a000016
- Simpson, G., & Clifton, J. (2016). Assessing postgraduate student perceptions and measures of learning in a peer review feedback process. *Assessment & Evaluation in Higher Education*, 41(4), 501-14. doi:10.1080/02602938.2015.1026874

- Sluijsmans, D. M. A., Brand-Gruwel, S., & vanMerriënboer, J. J. G. (2002). Peer assessment training in teacher education: effects on performance and perceptions. *Assessment and Evaluation in Higher Education*, 27(5), 443-454.
<https://doi-org.huaryu.kl.oakland.edu/10.1080/0260293022000009311>
- Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing*, 1(3), 217-233. doi:10.1016/1060-3743(92)90004-9
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
<https://doi-org.huaryu.kl.oakland.edu/10.1093/applin/16.3.371>
- Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). 9. peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, 22, 171-185.
doi:10.1017/S0267190502000090
- Tang, G. M., & Tithecott, J. (1999). Peer response in ESL writing. *TESL Canada Journal*, 16(2), 20-38. <https://doi.org/10.18806/tesl.v16i2.716>
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. Retrieved from
<http://journals.sagepub.com.huaryu.kl.oakland.edu/doi/pdf/10.3102/00346543068003249>
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education*, 25(2), 149-169. doi:10.1080/713611428
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27.
doi:10.1080/00405840802577569
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170.
doi:10.1016/S1060-3743(00)00022-9

- Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, 21(2), 217-235.
doi:10.1016/j.compcom.2004.02.003
- Van Steendam, E., Rijlaarsdam, G., Sercu, L., & Van den Bergh, H. (2010). The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. *Learning and Instruction*, 20(4), 316-327.
doi:10.1016/j.learninstruc.2009.08.009
- Villamil, O. S., & de Guerrero, M. C. M. (1996). Peer revision in the L2 classroom: Social-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing*, 5(1), 51-75.
doi:10.1016/S1060-3743(96)90015-6
- Vorobel, O., & Kim, D. (2014). Focusing on content: Discourse in L2 peer review groups. *TESOL Journal*, 5(4), 698-720. doi:10.1002/tesj.126
- Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S., & Kozulin, A. (1986). *Thought and language* (Translation newly rev. and ed.). Cambridge, Mass: MIT Press.
- Wells, C. G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179-200.
doi:10.1016/j.jslw.2006.09.004
- Yu, S., & Lee, I. (2014). An analysis of Chinese EFL students' use of first and second language in peer feedback of L2 writing. *System: An International Journal of Educational Technology and Applied Linguistics*, 47, 28-38.
doi:10.1016/j.system.2014.08.007
- Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly*, 16(2), 195-209. doi: 10.2307/3586792
- Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, 19(1), 79-101.
doi: 10.2307/3586773

- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166-183. doi:10.3102/0002831207312909
- Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing*, 4(3), 209-222. doi:10.1016/1060-3743(95)90010-1
- Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing*, 15(1), 3-17. doi:10.1016/j.asw.2010.01.002
- Zhao, H. (2014). Investigating teacher-supported peer assessment for EFL writing. *ELT Journal*, 68(2), 155. doi:10.1093/elt/cct068
- Zhu, W. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication*, 12(4), 492-528. doi:10.1177/0741088395012004004
- Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, 10(4), 251-276. doi:10.1016/S1060-3743(01)00043-1
- Zhu, W., & Mitchell, D. A. (2012). Participation in peer response as activity: An examination of peer response stances from an activity theory perspective. *TESOL Quarterly*, 46(2), 362-386. doi: 10.1002/tesq.22