

**Peer Feedback and Its Effects on Chinese  
English-as-a-Foreign-Language (EFL) Learners' Text  
Revisions and Writing Performance**

**Zhiqing Xu**

A thesis submitted in fulfillment of the requirements for the Degree of Doctor of  
Philosophy (PhD) in Education (Applied Linguistics and TESOL)

The University of Auckland

Auckland, New Zealand

2020

## **Abstract**

Peer feedback, an activity in which learners read others' drafts and make comments on them orally and/or in written mode, is widely employed in writing instruction. Extensive studies investigating its effects have been conducted. However, several issues remain under-researched, including long-term effects of peer feedback on writing performance, learners' processing of peer feedback in revision, and learners' perceptions of peer feedback. This thesis reports on a study that addresses these research gaps.

This study was conducted during a semester-long College English course at a Chinese university. It was comprised of three parts. Part One investigated the effects of a rigorous peer feedback intervention on text revisions and writing performance. The treatment group participants engaged in six writing tasks with peer feedback activities while the comparison group participants wrote on the same topics and received the conventional collective feedback from their teacher to the whole class. Their two drafts of one writing task were compared to examine effects of the intervention on text revisions and their compositions in the pre-, post-, and delayed post-tests were compared to investigate its effects on writing performance. Part Two was a case study in which think-aloud protocols were analyzed to explore how students processed peer feedback when writing their second drafts. As a final part, Part Three employed questionnaire and interviews to explore students' perceptions of the intervention, specifically any perceived changes in their writing performance and their perceived helpfulness of specific components within the intervention.

Results reveal that the intervention had positive effects on students' text revisions. The treatment group participants made significant improvements between drafts in the

overall text quality, content, organization, accuracy, and syntactic complexity, although not in lexical complexity or fluency. In addition, their second drafts were better in the overall text quality, content and organization than those written by the comparison group participants. The analysis of the think-aloud protocols showed that students utilized various cognitive operations and consequently two different approaches to process peer feedback when writing their second drafts. The intervention was also found to have positive effects on students' writing performance. The treatment group participants made significant improvements in the overall text quality, content, organization, accuracy in the post-test and they retained the improvements twelve weeks after the intervention. The treatment group participants outperformed their counterparts receiving collective feedback in the overall text quality and organization in the post- and delayed post-tests, and in accuracy only in the delayed post-test. The data from the questionnaire and interviews corroborated positive effects of the intervention on writing performance and suggested that there were other areas that the intervention might have positive effects on. These findings were discussed as to how specific components in the intervention as well as students' efforts in processing peer feedback contributed to improvements in text revisions and writing performance.

This study extends the existing literature on peer feedback in several ways. Theoretically, it has examined long-term effects of peer feedback on writing performance, adding strong evidence to the long-standing debate about the effects of peer feedback. The findings about cognitive operations and approaches in processing feedback suggest adaptations to Hayes' (1996) revision model. Pedagogically, this study provides L2 writing practitioners with practical insights into writing instruction in relation to teachers' incorporating peer feedback and helping students improve writing proficiency.

## Acknowledgements

As my four-year “PhD marathon” is drawing to a close, I would like to express my sincere gratitude to the support that I received from various parties, including my supervisors, my participants, my family and my employer.

I am deeply indebted to my main supervisor Professor Lawrence Jun Zhang in many ways. He offered me tremendous help and continuous support throughout my PhD journey. Under his supervision, I have grown from a frontline teacher to an academic researcher who takes the initiative to reflect upon and optimize my daily teaching practices. His patience with me is also highly appreciated because he allowed me to explore and learn at my own pace. Moreover, he has set up a community comprising more than 20 PhD students at different stages and promoted collaboration and support within the community, where I really experienced the meaning of what a “community of practice” stood for. I feel fortunate to be a member within the community.

I would like to extend my special thanks to my co-supervisor Professor Judy M. Parr. Her meticulous and critical comments on my thesis proposal and thesis draft helped me shape my research and provided me with valuable insights. Meanwhile, I thank her for her unrelenting encouragement and unfailing faith in me when I encountered difficulty and challenges in writing my thesis.

My gratitude goes to the participating teachers and students. Without their willingness to get involved in my research project, the thesis would not have been possible. Their collaboration and trust enabled me to gain valuable insights into their teaching and learning.

I shall always be thankful to my family. They respected my decision to study overseas and stood by me when in need. My thanks go to my parents and parents-in-law for traveling a long distance from China to New Zealand and helping take care of my son. My special thanks go to my husband Zhi Li for his love, understanding, and support.

This research was generously supported by a full PhD scholarship that the University of Auckland and the China Scholarship Council jointly awarded me, which freed me up from all financial concerns. I am genuinely indebted to them for their generosity. I also acknowledge my appreciation to my employer, Jilin University, for kindly granting me four years of research and study leave, which made it possible for me to complete my study.

# Contents

<b>Abstract.....</b>	<b>i</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>Contents.....</b>	<b>v</b>
<b>List of Tables .....</b>	<b>xii</b>
<b>List of Figures.....</b>	<b>xvi</b>
<b>List of Abbreviations .....</b>	<b>xvii</b>
<b>Chapter 1 Introduction.....</b>	<b>1</b>
1.1 Chapter overview .....	1
1.2 The research context.....	1
1.2.1 English as a foreign language (EFL) teaching in China .....	1
1.2.2 College English course and writing instruction within the course .....	2
1.3 Definition of key terms .....	4
1.3.1 Peer feedback .....	4
1.3.2 Collective feedback.....	5
1.3.3 Revision .....	6
1.4 Key issues and gaps in research on peer feedback.....	6
1.5 Research purposes and major research questions .....	8
1.6 Significance of the study .....	8
1.7 Outline of the thesis.....	9
<b>Chapter 2 Literature Review .....</b>	<b>11</b>

2.1	Chapter overview .....	11
2.2	Peer feedback in the process approach to writing .....	11
2.2.1	Process approach to writing .....	11
2.2.2	Establishment of peer feedback research .....	14
2.3	Effects of peer feedback on text revisions and writing performance .....	16
2.3.1	Studies on effects of peer feedback on text revisions .....	16
2.3.2	Studies on effects of peer feedback on writing performance.....	27
2.3.3	Training in peer feedback studies .....	35
2.3.4	Summary of findings on effects of peer feedback .....	37
2.4	Processing of peer feedback in revision .....	39
2.4.1	Major models of revision .....	40
2.4.2	Empirical findings of L2 learners' processing of feedback in revision ..	48
2.5	Student perceptions of peer feedback.....	55
2.5.1	L1 learners' perceptions of peer feedback .....	55
2.5.2	L2 learners' perceptions of peer feedback .....	56
2.5.3	Summary of findings on students' perceptions of peer feedback .....	59
2.6	Conceptual framework for the present study .....	60
2.7	Chapter summary .....	64
	<b>Chapter 3 Methodology .....</b>	<b>65</b>
3.1	Chapter overview .....	65
3.2	Overview of research design .....	65

3.2.1	Research paradigms and approaches .....	65
3.2.2	Research design .....	66
3.3	Instructional context.....	69
3.4	Preparatory stage: Instrument and procedure trial and modification .....	70
3.4.1	Instrument refinement.....	70
3.4.2	Intervention procedure trial.....	71
3.4.3	Some changes to the instruments and procedures.....	73
3.5	Procedures for the study.....	74
3.5.1	Pre-test, post-test, and delayed post-test weeks .....	75
3.5.2	Peer feedback training.....	75
3.5.3	Six writing tasks with peer feedback activities.....	78
3.5.4	Three independent writing tasks .....	79
3.6	Part One: Effects of peer feedback on text revisions and writing performance.. .....	80
3.6.1	Participants.....	81
3.6.2	Data collection .....	82
3.6.3	Data analysis .....	84
3.7	Part Two: Processing of peer feedback in revision .....	90
3.7.1	Participants.....	91
3.7.2	Data collection .....	91
3.7.3	Data analysis .....	95
3.8	Part Three: Perceptions of peer feedback.....	102



3.8.1	Data collection .....	103
3.8.2	Data analysis .....	105
3.9	Ethical considerations .....	106
3.9.1	Informed consent .....	106
3.9.2	Right to withdraw from participation .....	106
3.9.3	Confidentiality .....	106
<b>Chapter 4</b>	<b>Effects of Peer Feedback Intervention .....</b>	<b>107</b>
4.1	Chapter overview .....	107
4.2	Effects of peer feedback on text revisions .....	108
4.2.1	Comparison of baseline conditions of the two groups.....	108
4.2.2	Changes in writing scores between drafts.....	110
4.2.3	Changes in CAF indexes between drafts .....	113
4.3	Effects of peer feedback on writing performance .....	116
4.3.1	Comparison of baseline conditions of the two groups.....	116
4.3.2	Changes in writing scores .....	117
4.3.3	Changes in CAF indexes.....	127
4.4	Chapter summary .....	138
<b>Chapter 5</b>	<b>Case Study Findings .....</b>	<b>142</b>
5.1	Chapter overview .....	142
5.2	Received peer feedback points.....	143
5.3	Processing of peer feedback in revision.....	144

5.3.1	Cognitive operations .....	145
5.3.2	Approaches .....	163
5.4	Chapter summary .....	175
<b>Chapter 6 Student Perceptions of Peer Feedback.....</b>		<b>176</b>
6.1	Chapter overview .....	176
6.2	Perceived changes in writing performance .....	177
6.2.1	Content and organization .....	178
6.2.2	Language use .....	184
6.3	Perceived usefulness of specific components in the intervention .....	190
6.3.1	General training and focused training.....	191
6.3.2	Giving feedback and receiving feedback.....	196
6.3.3	Paired discussions .....	203
6.3.4	PF forms.....	207
6.4	Chapter summary .....	211
<b>Chapter 7 Discussion.....</b>		<b>214</b>
7.1	Chapter overview .....	214
7.2	Effects of peer feedback on text revisions .....	214
7.2.1	Overall quality .....	214
7.2.2	Content and organization qualities.....	216
7.2.3	Language use .....	219
7.3	Effects of peer feedback on writing performance .....	223

7.3.1	Overall quality .....	223
7.3.2	Content and organization .....	224
7.3.3	Language use .....	227
7.3.4	Retention of effects for peer feedback on writing performance .....	231
7.4	Effectiveness of peer feedback intervention .....	233
7.4.1	Reciprocal peer feedback activities .....	233
7.4.2	Paired discussions .....	235
7.4.3	Various forms of teacher support.....	236
7.5	Students' efforts in processing peer feedback.....	238
7.5.1	Cognitive operations and their efficacy .....	239
7.5.2	Two approaches and their efficacy .....	242
7.6	Issues in peer feedback intervention .....	244
7.7	Further development of Hayes' (1996) revision model .....	247
7.8	Chapter summary .....	251
<b>Chapter 8</b>	<b>Conclusions, Implications, and Recommendations.....</b>	<b>252</b>
8.1	Chapter overview .....	252
8.2	Summary of major findings.....	252
8.2.1	Effects of peer feedback on text revisions .....	252
8.2.2	Effects of peer feedback on writing performance.....	253
8.2.3	Processing of peer feedback in revision.....	254
8.2.4	Perceptions of peer feedback intervention .....	254

8.3	Contributions and implications .....	254
8.3.1	Theoretical contributions .....	255
8.3.2	Methodological contributions .....	255
8.3.3	Pedagogical implications .....	256
8.4	Limitations and recommendations for future research.....	257
8.5	Conclusion.....	258
<b>References.....</b>		<b>260</b>
<b>Appendix A: Scoring Rubric Adapted from Jacobs et al. (1981)'s ESL Composition Profile.....</b>		<b>283</b>
<b>Appendix B: Writing Prompts for the Six Writing tasks during the Intervention Period.....</b>		<b>285</b>
<b>Appendix C: Peer Feedback Form for Writing Task 6.....</b>		<b>287</b>
<b>Appendix D: Writing Prompts for Three Independent Writing Tasks during the Post-intervention Period.....</b>		<b>294</b>
<b>Appendix E: Writing Prompts for and Pre-, Post-, and Delayed post-tests.....</b>		<b>295</b>
<b>Appendix F: Error Coding Scheme.....</b>		<b>296</b>
<b>Appendix G: Coding of revision episodes &amp; cognitive operations in PF triggered revision episodes.....</b>		<b>298</b>
<b>Appendix H: The Perceptions of Peer Feedback Questionnaire .....</b>		<b>301</b>
<b>Appendix I: Questions for the Semi-structured Interviews.....</b>		<b>306</b>
<b>Appendix J: Ethics Forms.....</b>		<b>307</b>

## List of Tables

Table 3.1 An overview of the research design.....	68
Table 3.2 Procedures for the study .....	74
Table 3.3 Stages and activities for the general training session .....	75
Table 3.4 Topics in the six focused peer training sessions .....	77
Table 3.5 Participants in the quasi-experimental study .....	82
Table 3.6 Compositions collected for the quasi-experimental study .....	83
Table 3.7 Coding scheme for feedback items .....	96
Table 3.8 Coding scheme for cognitive operations in peer-feedback-triggered revision episodes.....	100
Table 4.1 Descriptive statistics and one-way ANOVA analysis of scores for Draft 1 of Writing Task 6 .....	109
Table 4.2 Descriptive statistics and one-way ANOVA analysis of CAF indexes for Draft 1 of Writing Task 6 .....	110
Table 4.3 Group means and standard deviations of the scores for Draft 1 and Draft 2 of Writing Task 6 .....	111
Table 4.4 Within-group differences in scores between drafts of Writing Task 6.....	112
Table 4.5 Group means and standard deviations of CAF indexes for Draft 1 and Draft 2 of Writing Task 6 .....	114
Table 4.6 Within-group differences in CAF indexes between drafts of Writing Task 6 .....	115

Table 4.7 Descriptive statistics and one-way ANOVA analysis of scores in the pre-test .....	117
Table 4.8 Descriptive statistics and one-way ANOVA analysis of CAF indexes in the pre-test.....	117
Table 4.9 Group means and standard deviations of overall scores in the pre-, post-, and delayed post-tests .....	118
Table 4.10 Effect of time and group condition on overall scores .....	119
Table 4.11 Within-group comparisons in overall scores .....	120
Table 4.12 Between-group comparisons in overall scores .....	121
Table 4.13 Group means and standard deviations of content scores in the pre-, post-, and delayed post-tests .....	122
Table 4.14 Effect of time and group condition on content scores .....	123
Table 4.15 Within-group comparisons in content scores.....	123
Table 4.16 Group means and standard deviations of organization scores in the pre-, post-, and delayed post-tests .....	124
Table 4.17 Effect of time and group condition on organization scores .....	125
Table 4.18 Within-group comparisons in organization scores .....	126
Table 4.19 Between-group comparisons in organization scores .....	127
Table 4.20 Group means and standard deviations of EP100W in the pre-, post-, and delayed post-tests .....	128
Table 4.21 Effect of time and group condition on EP100W.....	129
Table 4.22 Within-group comparisons in EP100W .....	130

Table 4.23 Between-group comparisons in EP100W .....	130
Table 4.24 Group means and standard deviations of MLT in the pre-, post-, and delayed post-tests .....	131
Table 4.25 Effect of time and group condition on MLT.....	133
Table 4.26 Group means and standard deviations of MSTTR_50 in the pre-, post-, and delayed post-tests .....	134
Table 4.27 Effect of time and group condition on MSTTR_50.....	135
Table 4.28 Group means and standard deviations for WPT in the pre-, post-, and delayed post-tests .....	136
Table 4.29 Within-group comparisons in WPT .....	137
Table 4.30 Between-group comparisons in WPT .....	138
Table 4.31 Changes in writing scores and CAF indexes in Writing Task 6.....	139
Table 4.32 Changes in writing scores and CAF indexes over time .....	141
Table 5.1 Case study participants' demographic information .....	143
Table 5.2 Peer feedback items received on Draft 1 of Writing Task 6.....	144
Table 5.3 Peer feedback items received and attended to in revision .....	145
Table 5.4 The cognitive operations and their frequencies in the 94 peer-feedback-triggered revision episodes .....	146
Table 5.5 The direct approach and elaborated approach .....	164
Table 6.1 Treatment group participants' perceptions of their improvements in writing performance .....	177
Table 6.2 Students' perceived changes in content .....	178

Table 6.3 Student' perceived changes in organization .....	178
Table 6.4 Students' perceived changes in grammar .....	184
Table 6.5 Treatment group participants' perceptions of specific components in the intervention .....	191
Table 6.6 Students' perceptions of the general training session.....	192
Table 6.7 Students' perceptions of focused training sessions.....	194
Table 6.8 Students' perceptions of giving feedback.....	197
Table 6.9 Students' perceptions of receiving feedback .....	197
Table 6.10 Students' perceptions of paired discussions .....	204
Table 6.11 Students' perceptions of using peer feedback forms .....	208
Table 7.1 Changes between pre-test and post-test and between pre-test and delayed post-test.....	232



## List of Figures

Figure 2.1 The Flower and Hayes' (1981) cognitive process writing model .....	12
Figure 2.2 W. Wang and Wen's (2002) L2 writing model .....	14
Figure 2.3 The Hayes et al.'s (1987) cognitive process model of revision .....	41
Figure 2.4 The Hayes' (1996) cognitive process model of revision.....	44
Figure 2.5 Butterfield et al.'s (1996) cognitive-metacognitive model of revision .....	45
Figure 2.6 Rajoo's (2009) model of thought processes on teacher feedback .....	52
Figure 3.1 Convergent design .....	66
Figure 4.1 Mean overall scores in the pre-, post-, and delayed post-tests .....	119
Figure 4.2 Mean content scores in the pre-, post-, and delayed post-tests .....	122
Figure 4.3 Mean organization scores in the pre-, post-, and delayed post-tests .....	125
Figure 4.4 The mean number of EP100W in the pre-, post-, and delayed post-tests .	128
Figure 4.5 The MLT in the pre-, post-, and delayed post-tests.....	132
Figure 4.6 The MSTTR_50 in the pre-, post-, and delayed post-tests.....	134
Figure 4.7 The WPT in the pre-, post-, and delayed post-tests.....	136
Figure 7.1 The adapted Hayes' (1996) cognitive process model of revision .....	248

## List of Abbreviations

EFL	English as a foreign language
ESL	English as a second language
ESP	English for specific purposes
SLA	Second language acquisition
L1	English as a first language
L2	English as a second language
CET	College English Test
TEM	Test for English majors
PF	Peer feedback
CAF	Complexity, accuracy and fluency
EP100W	Errors per 100 words
MLT	Mean length of T-units
MSTTR_50	Mean segmental type-token ration per 50 words
WPT	Words per text

# Chapter 1 Introduction

## 1.1 Chapter overview

This chapter first presents the research context by introducing English teaching and learning in the Chinese education system and the English writing instruction for non-English majors in Chinese tertiary institutes. Next, it presents key terms in the thesis and highlights gaps in peer feedback literature. What follows are the objectives and research questions, and the significance of the present study. This chapter ends with an outline of the thesis.

## 1.2 The research context

### 1.2.1 English as a foreign language (EFL) teaching in China

English, with its established status of a lingua franca in many domains worldwide, is widely recognized as a critical second, or foreign language in China (Cheng & Wei, 2019; Gil, 2016; Woodrow, 2011). In the past several decades, the Chinese government and the Ministry of Education have put an emphasis on English education and established an EFL teaching system covering primary, secondary and tertiary education.

EFL education starts from Year 1 in primary school (6 years in total). Primary school students have an average of three sessions of English course per week (each session for 45 minutes), and they experience a communicative teaching style during this stage.

In secondary education, three years of junior secondary education and three years of senior secondary education, English is a compulsory course within the curriculum. Students take an average of five sessions each week and are required to pass an English

examination at the end of their junior secondary schooling and again at the end of their senior secondary schooling. Their scores in these two examinations serve as essential criteria for their entrance to senior secondary schools and tertiary institutes. EFL education at this stage is grammar-focused and exam-oriented (He & Teng, 2019; Wang & Gao, 2008).

Tertiary education consists of undergraduate, postgraduate, and doctoral programs with English as a compulsory course for all these programs. Undergraduates, who major in a discipline other than English, take English courses which focus on the five basic language skills (listening, reading, translating, writing and speaking). If they want to apply for postgraduate programs and then doctoral programs, they are required to sit the National English Test for Postgraduate Entrance Examination, and the Graduate Admission Test of English for Doctoral Candidates respectively. Most non-English majors also choose to sit College English Test (CET), a national standardized test that evaluates their English proficiency from five aspects, including listening, reading, translating, writing and speaking. English majors have a different program, which encompasses various courses, such as linguistics, literature, translation and academic writing; they take the Test for English Majors (TEM) to obtain evidence of their language proficiency.

### **1.2.2 College English course and writing instruction within the course**

The English course for non-English majors is referred to as College English. It is an integrated course, targeting “students’ ability of using English, their cross-cultural awareness and communicative ability...[as well as their] learner autonomy and comprehensive cultural literacy” (Cai, 2017, p. 117). All non-English majors are required to take the course for four semesters, with each semester lasting for 17 weeks

on average and each week having two sessions (90 minutes each). The course is typically delivered in a classroom with 45 to 60 students.

Over the past three decades, College English teaching has undergone massive reforms. Firstly, College English teaching has been implemented through compulsory courses, elective courses, or a combination of both, arranged flexibly within the four years of undergraduate study (X. Zhou & Zhan, 2016). Secondly, various approaches have been incorporated in teaching, such as communicative language teaching, task-based and project-based instruction, collaborative learning, and inquiry-based approaches. Thirdly, with the rapid development and application of education technologies, multi-media and information technologies have revolutionized the traditional College English course, improving teaching efficiency as well as facilitating individualized and autonomous learning. Flipped learning, Massive Open Online Courses (MOOC), and other rich online resources complement traditional College English teaching and ease the constraints of time and space for teaching (S. Wang, 2016).

Writing, a required component in the curriculum, however, occupies a “Cinderella” position in College English teaching in China, and also in many parts of the world where English is taught as a second or foreign language (L. J. Zhang, 2016, p. 207). In contrast to its status as an independent course for English majors in China and for university students in English-speaking countries, writing is typically taught through the College English course, with emphasis given to linguistic accuracy and rote learning (Huang & Zhang, 2020; Ng & Cheung, 2017). When the CET is imminent, EFL teachers allocate some time for writing practice and analysis of model essays (You, 2004; L. J. Zhang, Liu, & Liu, 2018). Also, the large class size prevents teachers from offering feedback in a timely fashion (Huang, 2020).

Under the influence of reforms in College English teaching, writing instruction has also undergone several changes. Firstly, approaches imported from English as a first language (L1) writing, such as process approach and genre approach, have been incorporated in the traditional writing instruction. Information technology and online resources have also been utilized in writing instruction. For example, Pigaiwang, the most popular commercial automated writing evaluation system in China, is now being used by “over 1,000 schools/universities, 40,000 English teachers and 4.5 million English learners” in assessing “over 180 million compositions” (Bai & Hu, 2017, p. 70). Furthermore, some institutes have started to offer writing courses as electives, preparing students for their academic needs.

### **1.3 Definition of key terms**

To facilitate understanding of the following thesis, I provide definitions for three key terms used in the thesis: Peer feedback, collective feedback and revision.

#### **1.3.1 Peer feedback**

Peer feedback in the present study is defined as an activity, in which learners “provide and receive feedback on their peers’ writing in the written and/or oral mode in pairs or small groups” (Yu & Lee, 2016, p. 461). This definition accentuates two features of feedback congruent with the present study: the reciprocal nature and the multiple modes of peer feedback. In the present study, peer feedback was implemented in the revising stage when students finished their first drafts and conducted in pairs as a reciprocal activity; it involved written feedback and oral feedback exchanged in discussions.

In the field of writing research, there are many synonyms for peer feedback, including peer review, peer response, peer editing and peer assessment. Although the differences

among these terms are not essential for the present study, I still feel a need to clarify the differences and explain my decision to use peer feedback rather than other terms. Among the four terms, peer review often denotes provision of written feedback (e.g., Hu & Lam, 2010; Min, 2006) and so may be confused with reviews of academic papers submitted for publication; peer response is often used in studies with a focus on oral feedback (e.g., Lockhart & Ng, 1995; Nelson & Carson, 1998); peer editing conveys a focus on language errors (e.g., Diab, 2010); and peer assessment typically involves “rating of writing quality” in addition to provision of feedback (Chang, 2016, p. 82). As the present study focuses on written feedback and, at the same time, involves oral feedback as a supplement, these four terms do not convey such a focus. I therefore decided to use peer feedback throughout the thesis. The other four terms will only be occasionally used, as appropriate, when other studies are cited in the literature review and discussion chapters.

### 1.3.2 **Collective feedback**

Collective feedback, an approach of teacher feedback, refers to a teacher’s practice of evaluating a sample of compositions and addressing some common problems to the whole class orally or through a written feedback form (Gallien & Oomen-Early, 2008). Although this approach may not “address personal needs and provide opportunities for personal interaction as individual feedback”, it has the advantage of being timely and satisfying the need for frequent teacher feedback (Gielen, Tops, Dochy, Onghena & Smeets, 2010). In this study, the teacher participant, for each writing task, evaluated 1/2 to 1/3 of compositions collected from the comparison group participants, summarized common problems, and finally presented them to the whole class orally together with the use of PowerPoint slides. In a personal communication with the teacher participant, she explained that such a teacher feedback approach was widely used in the university

where data for my research were collected, because it satisfied the need of responding to a large number of students' compositions within a limited period.

### **1.3.3 Revision**

Revision has two referents in writing research. When used as an uncountable noun, it refers to the mental process of making changes to a piece of writing; when used as a countable noun and mostly in plural form, revisions refer to physical changes in a text resulting from revision (Allal & Chanquoy, 2004). In the present study, to make a distinction between the two referents, I will use "revision" and "text revisions" to refer to the process and physical changes between the two drafts, respectively. As a process, revision is composed of revision episodes in which writers address problems in their first drafts, triggered by peer feedback or other factors. Since one focus of this study was on students' processing of peer feedback, only peer-feedback-triggered (PF-triggered) revision episodes were analyzed in terms of cognitive operations. Also, text revisions were examined in terms of changes in text quality between the drafts.

## **1.4 Key issues and gaps in research on peer feedback**

Peer feedback, as a major activity in writing instruction associated with social, cognitive, affective, and linguistic benefits, has received sustained attention from researchers and practitioners. By engaging in peer feedback activities, students gain multiple perspectives on their writing, experiment with and acquire a variety of language and writing skills, and develop audience awareness and learner autonomy (Min, 2005; Lundstrom & Baker, 2009; Tsui & Ng, 2000; L. J. Zhang & Cheng, 2020, 2021).

A review of L2 studies on peer feedback, however, reveals several research gaps. Firstly, the majority of existing L2 studies have focused on the effects of peer feedback on text



revisions. L2 studies, by using several revision measures and writing quality measures, have produced mixed results (e.g., Lam, 2013; Tsui & Ng, 2010; Zhao, 2010). Also, there have been methodological problems in these studies, such as the lack of a comparison group, the use of more than one feedback form with the same group of participants, and the inclusion of a limited number of writing tasks with peer feedback activities. There is a need to address these methodological issues in peer feedback studies. Furthermore, only a few studies have investigated the impact of peer feedback on writing development, (e.g., Diab, 2016; Ruegg, 2015a). Peer feedback studies which trace writing development over an extended period of time are therefore needed (Yu & Lee, 2016).

Another research gap lies in the limited investigation into how learners process peer feedback in revision. Most peer feedback studies, whether in L1 or L2 context, have relied on product data (compositions). Very few studies have explored students' processing of peer feedback in revision; no peer feedback study, up to date, has employed the think-aloud method (see L. J. Zhang & Cheng, 2021). In the present study, think-aloud protocols were used to provide process data to investigate the cognitive operations that students employ in processing peer feedback in revision.

The third research gap is the insufficient knowledge of learners' perceptions of peer feedback. Existing studies (e.g., van den Berg, Admiraal & Pilot, 2006; J. Zhou, Zheng & Tai, 2020) have investigated learners' general attitudes towards peer feedback; few studies have examined learners' perceived changes in different areas of writing and their perceptions of specific elements within the peer feedback intervention. Since such findings can help interpret the efficacy of peer feedback and possibly reveal some issues

affecting its efficacy, data on learners' perceptions needs to be incorporated in peer feedback studies.

### **1.5 Research purposes and major research questions**

This study investigates effects of peer feedback on students' text revisions and writing performance and addresses the following four major questions:

- (1) How does peer feedback influence Chinese EFL learners' text revisions?
- (2) How does peer feedback influence Chinese EFL learners' writing performance?
- (3) How do Chinese EFL learners process peer feedback when writing their second drafts?
- (4) How do Chinese EFL learners perceive the effectiveness of the peer feedback intervention?

### **1.6 Significance of the study**

Theoretically, this study extends the existing literature on peer feedback and contributes to our understanding of revision. The present study, by including a delayed post-test in the research design, examined the long-term impact of peer feedback on writing performance, adding strong evidence to the on-going debate on the effects of peer feedback. The findings regarding the cognitive operations and approaches in processing peer feedback, further developed Hayes' (1996) revision model.

Methodologically, this study features novelty in its research design. In terms of duration of the intervention, it involved 13 weeks with six writing tasks with peer feedback activities. Storch (2010) viewed the duration of treatment as a major concern in studies

on feedback and notes the need for sustained feedback; Ruegg (2020) suggested that feedback studies be conducted over a semester-long course and implement no less than four feedback iterations. As progress in writing occurs over an extended period of time, the administration of one-off, or very few, treatments lacks theoretical and pedagogical validity, as some scholars posited (e.g., Storch, 2010). In addition, this study measured the quality of writing using writing scores as well as complexity, accuracy and fluency (CAF) indexes, which reflected learners' writing performance dynamically. Also, this study employed multiple instruments (writing tests, think-aloud protocols, questionnaire, and interview) and incorporated multiple datasets (product data, process data and perception data). Research findings from this study therefore allow a more confident interpretation and thorough understanding of the effects of peer feedback.

Thirdly, this study informs feedback practices for the teaching of writing in the Chinese EFL context as well as, potentially, other EFL contexts. Since the study reveals how and in what aspect peer feedback influenced Chinese EFL learners' text revisions and writing performance, the findings will provide implications for designing and implementing peer feedback activities in writing classrooms.

## **1.7 Outline of the thesis**

The thesis consists of seven chapters. As an introduction, Chapter 1 starts with an overview of the context in which the research project was conducted. It then summarizes research gaps in peer feedback research, and identifies research purposes, research questions, and significance of the study. Chapter 2 reviews previous literature in three related areas: effects of peer feedback on text revisions and writing performance; processing of feedback in revision; and learners' perceptions of peer feedback. The chapter concludes by presenting the theoretical framework for this study. Chapter 3,

which describes the research methodology for this study, begins with an overview of its research design and then introduces its instructional context, preparatory stage, and procedures. It then introduces the three parts of the study including participants, instruments used for data collection and methods and procedures for data analysis. The next three chapters, Chapter 4, Chapter 5 and Chapter 6 report the findings of this study: Chapter 4 reports on the effects of peer feedback on text revisions and writing performance, as measured by writing scores and CAF indexes; Chapter 5 describes how students processed peer feedback in revision; and Chapter 6 presents students' perceptions of peer feedback. Chapter 7 discusses the results reported in the three finding chapters. The final chapter, Chapter 8, presents a summary of this study followed by a consideration of theoretical and practical implications derived from the research findings. The thesis concludes with a discussion of the limitations of this study and suggestions for future studies.

## **Chapter 2 Literature Review**

### **2.1 Chapter overview**

This chapter reviews relevant L1 and L2 peer feedback studies as well as theories about feedback and revision. First, it reviews literature on the process approach to writing, the role of peer feedback in the process-based writing instruction, as well as theories that inform peer feedback. Empirical L1 and L2 studies focusing on effects of peer feedback, learners' processing of feedback in revision and learners' perceptions of peer feedback are then reviewed. The final section of the review explicates the conceptual framework for the present study. The chapter ends with a summary outlining the major findings and research gaps in the literature.

### **2.2 Peer feedback in the process approach to writing**

#### **2.2.1 Process approach to writing**

Since the 1970s, L1 and L2 writing researchers have shifted their focus from written products to writing processes. At the same time, various cognitive models of composing have been proposed which have contributed to the establishment of process approach to writing.

The early form of process approach viewed writing as a linear process, comprised of prewriting, composing, and revising. When working on a topic, writers move sequentially from prewriting in which they plan for organization and content of their texts, to composing in which they put their ideas on paper, and finally to revising in which they examine their written texts and make changes. This linear view of writing, however, was criticized for under-conceptualizing and oversimplifying the writing process (Emig, 1971).

The seminal work by Flower and Hayes (1981) had a revolutionary effect on the process approach. Writing began to be conceptualized as a recursive process involving planning, translating, reviewing, and monitoring (see Figure 2.1). Writers, rather than move sequentially through each activity, engage in more than one activity simultaneously and move back and forth between the activities. In writing courses, teachers emphasize the recursive nature of writing and help learners develop strategies of planning, drafting, revising and editing. They guide learners through the writing process by means of intervention activities and direct their attention to content and ideas in writing.

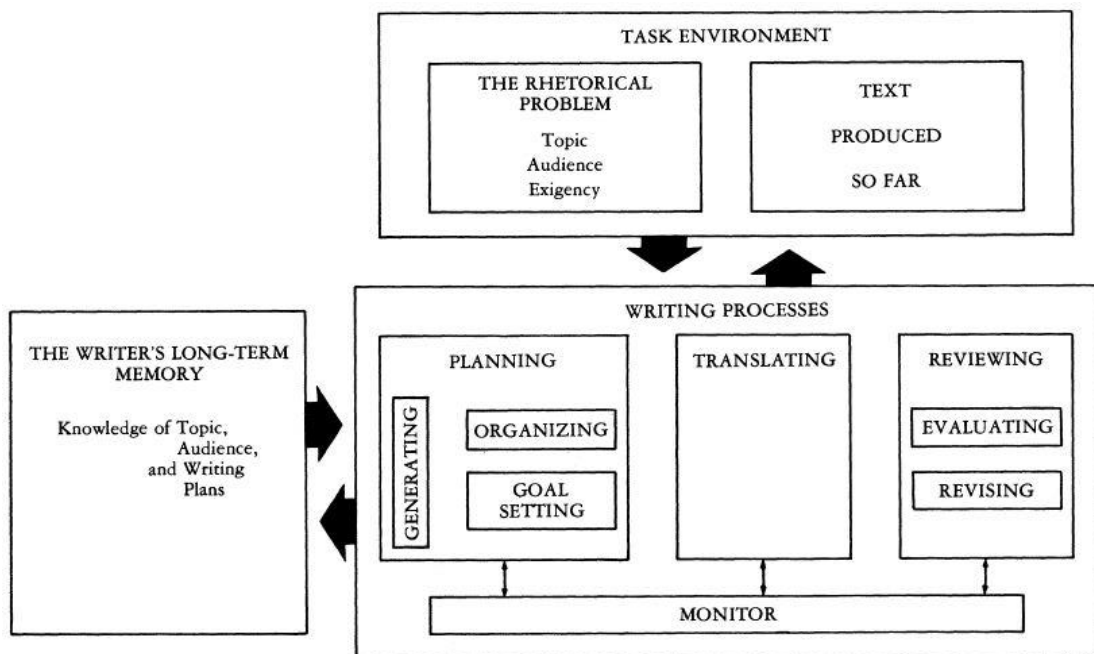


Figure 2.1 The Flower and Hayes' (1981) cognitive process writing model

Based on the think-aloud protocols from 16 Chinese EFL learners when composing, W. Wang and Wen (2002) proposed a L2 writing model (see Figure 2.2). The task environment and the writer's long-term memory were adopted from Flower and Hayes' (1981) writing model, while the composing processor was a modified component.

Specifically, the four sub-components of the writing process in the Flower and Hayes' (1981) writing model, planning, translating, reviewing and monitor, were replaced by five composing activities, namely task examining, idea generating, idea organizing, text generating, and process controlling. Bidirectional arrows were also added between these activities to visually present the recursive nature of writing as advocated by Flower and Hayes (1981). More importantly, W. Wang and Wen (2002) incorporated "the linguistic codes for the thoughts involved in each part and its subcomponent" in their L2 writing model (p. 241). As shown in Figure 2.2, both the input and output in the task environment are in L2. For the various activities in the composing processor, L2 learners are inclined to rely on L1 when monitoring their writing processes, and to use L2 when performing task-examining and text-generating activities. Of the sub-components in the writer's long-term memory, world knowledge and rhetorical knowledge are mainly in L1 whereas linguistic knowledge is mainly in L2.

The L2 writing model proposed by W. Wang and Wen (2002), in comparison with Flower and Hayes' (1981) cognitive process writing model, addressed the unique features of L2 writing, particularly the indispensable roles that L1 played in L2 writing. In addition, W. Wang and Wen (2002) found that L2 proficiency had a negative influence on writers' L1 use: "[L]ess proficient writers generate sentences through L1-to-L2 translation, while proficient writers generate text directly in L2" (p.240). As regards L1 use in L2 writing, other researchers cautioned against an over-reliance on L1 (Fujii, 2012; Kim, 2014). Having difficulty in organizing L2 words in proper syntactical structures, most low-proficiency L2 students were found to produce "substantially ungrammatical and awkwardly constructed" sentences (Fujii, 2012, p. 40).

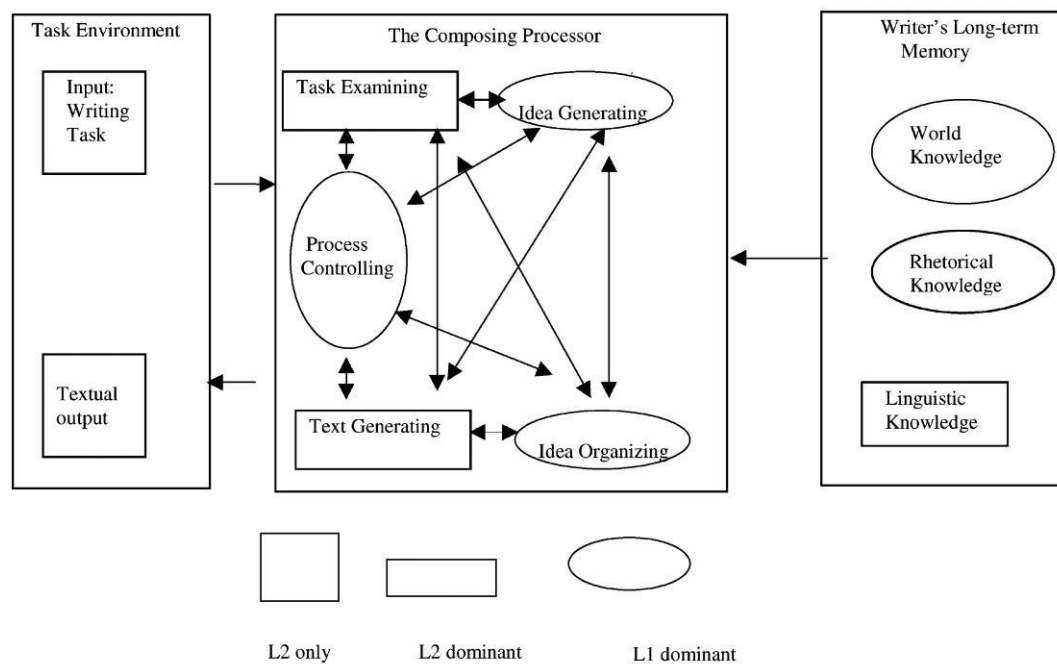


Figure 2.2 W. Wang and Wen's (2002) L2 writing model

Later, the process approach moved further towards the socio-cognitivism, as can be seen from researchers' consideration of more social factors (Atkinson, 2003a, 2003b; Casanave, 2003; Matsuda, 2003). For instance, Cumming, Busch and Zhou (2002) suggested that writing strategies be researched in reference to “the goals people have to motivate and guide their task performance as well as other essential aspects of these activity structures and the contexts in which they are embedded” (p.193); Atkinson (2003b) suggested reconceptualizing the concept of culture to identify social and cultural factors impacting L2 writing.

### 2.2.2 Establishment of peer feedback research

Peer feedback, a major component in the process-based writing instruction, is informed by the collaborative learning theory, the interactionist theory in second language acquisition (SLA), and social constructivism.



Collaborative learning theory contends that learning is socially constructed in that learners pool their resources and share the responsibility of learning (Trimbur, 1989). Peer feedback, a common pair/group activity in writing classrooms, creates a “facilitative socio-interactive environment in which L2 learners receive social support and scaffolding from peers” (Hu & Lam, 2010, p. 373). It is through collaboration and interaction with their peers that learners acquire linguistic knowledge as well as writing skills.

The interactionist theory, in SLA, maintains that interactions provide learners with opportunities to negotiate meaning and facilitate language acquisition (Long & Porter, 1985). In peer feedback activities, authentic interactions between the reader and the writer contribute to writing development in two ways. Firstly, the communication between the reader and the writer, particularly their negotiations in discussions, make the language input comprehensible. Secondly, peer feedback helps learners notice their problems and motivates them to revise their texts as output, serving as evidence for language acquisition.

Peer feedback is informed by social constructivism, which argues that higher forms of learning and cognitive development result from social interactions in which an individual works with peers within their Zone of Proximal Development (Vygotsky, 1978). In writing instruction, peer feedback constitutes a collaborative activity which helps learners achieve the transition from interpsychological functioning to intrapsychological functioning (Villamil & Guerrero, 2019).

With a strong theoretical basis as mentioned above, peer feedback constitutes a legitimate area in both L1 and L2 writing research, drawing the attention of researchers as well practitioners (Cao, Yu, & Huang, 2019; Gao, Schunn, & Yu, 2019). The

following three sections review peer feedback studies from three aspects: its effects on revision and writing performance, learners' processing of peer feedback in revision, and learners' perceptions of peer feedback.

## **2.3 Effects of peer feedback on text revisions and writing performance**

### **2.3.1 Studies on effects of peer feedback on text revisions**

As provision of feedback links naturally to making revisions in a subsequent draft, effects of peer feedback can be examined through revised drafts. L1 and L2 writing researchers, however, adopted different measures when investigating effects of peer feedback on text revisions. In the following subsections, L1 and L2 studies will be reviewed separately.

#### ***2.3.1.1 L1 studies on effects of peer feedback on text revisions***

Peer feedback, though commonly used in writing classrooms, has not been extensively researched in L1 settings (see Anson & Anson, 2017; Cho & MacArthur, 2011; Haswell, 2005; Thirakunkovit, 2019). A search for L1 studies on effects of peer feedback on text revisions has found only a few studies, conducted with learners at different school levels. Most of these studies were quasi-experimental, with one or more treatment groups and a comparison group; outcomes were reported in writing scores and revision measures.

Graner (1987) compared the impact of peer-editing and revision workshops on post-secondary students' text revisions. In this study, between the two drafts of each writing task, students in the peer-editing group shared their first drafts with other group members while those in the revision workshop group critiqued two anonymous compositions and participated in teacher-led discussions. Measured by an overall score, the researcher found that the two groups made similar improvement in their revised

drafts and concluded that the revision workshop was as effective as peer editing. Findings from this study, however, need to be interpreted with caution because several points in research design remain unclear. For instance, no mention was made to the size of peer editing group; whether peer editing activities were conducted during class hours with opportunities for discussions or whether training was provided to the peer editing group, was not explicated. These issues, to some extent, may have influenced the result.

Other L1 studies employed revision measures in addition to writing scores when examining the effects of peer feedback on text revisions. Olson (1990) compared the two drafts of a narrative task written by sixth-grade students in four conditions (receiving revision instruction and working with peer partners, working with peer partners only, receiving revision instruction only, and control group) and found the four groups differed in revision frequency, types and writing scores. The two groups having peer partners made fewer revisions than the other two groups; the revision instruction only group made content revisions more often while the other three groups made surface structure revisions and substitutions more often. As for writing scores, all the four groups made significant gains in surface structure quality scores (wording, syntax, usage, punctuation/capitalization, spelling, and appearance) between drafts; the peer partner only group made significant gains in rhetorical quality scores (audience considerations, voice, central figure, setting/background, overall organization, and theme/topic) between drafts. Because Olson (1990) attributed the revision instruction and peer partner group's lack of improvements in rhetorical quality to the focus of revision instruction being on accuracy and the already superior rhetorical quality for their first drafts, the findings from this study may suggest that peer feedback, whether used in conjunction with revision or used in isolation, helped young L1 students revise in narrative mode.

Similarly, Cho and Schunn (2007) and Cho and MacArthur (2010) examined the effects of three feedback forms on undergraduates' text revisions. Cho and Schunn (2007) included 28 students who were randomly assigned to three conditions, receiving feedback from a single peer, from a single expert, and from multiple peers (six anonymous peers). They reported that students receiving feedback from multiple peers made the greatest improvement in text quality between drafts and they significantly outperformed those in the other two conditions on their revised drafts. Moreover, they found that students, who received feedback from a single peer, made a slight improvement in text quality between drafts, but they performed similarly in text quality of their revised drafts as those receiving feedback from a single expert. Although the researchers attributed the unsatisfactory performance of students in the single expert condition to expert feedback being unhelpful and expert evaluation of the revised drafts being stringent, their findings suggest that multiple peer feedback was superior in helping students improve their drafts, and single peer feedback was at least as effective as single expert feedback for text revisions.

As a follow-up, Cho and MacArthur (2010) examined students' revisions in response to the three forms of feedback. Students in the multiple-peer group made more complex repairs (fixing points by deleting or changing existing points at the sentence or paragraph level) than those in the other two groups, and more extended content revisions (elaborating and clarifying an existing point) than those in the single expert group. Compared with previous studies, Cho and MacArthur (2010) made an important contribution by explaining the links between feedback and revisions, as well as between revisions and text quality, that is, non-directive feedback points (comments on a detail without suggesting a change) were predictive of complex repair revisions, and complex repair revisions were predictive of improved quality in the revised draft.

The two studies, Cho and Schunn (2007) and Cho and MacArthur (2010), taken together, provided strong evidence for, and an explanation of, the positive impact of multiple peer feedback on text revisions. The changes in writing scores reported in Cho and Schunn (2007), however, need to be scrutinized. Because the rubric used to evaluate text quality included three criteria, i.e., prose flow (sequencing of and transitions between points of argument), argument (quality of claims and support for arguments and counter-arguments), and insight (new ideas and information beyond the content of assigned materials), we can see that the writing scores reflected content and organization qualities, rather than the overall quality of the texts, for the two drafts.

In a study conducted in the secondary school setting, Lindgren (2005) considered language ability levels and genre differences when examining the effects of peer feedback on text revisions. With peer feedback, secondary students with low-L1-ability improved their draft quality for both the descriptive writing task and the argumentative writing task whereas those with high-L1-ability improved their draft quality for only the argumentative writing task. In addition, both low-L1-ability and high-L1-ability students made text-based and audience-oriented revisions. The results from this study, however, may not constitute strong evidence for the positive effects of peer feedback on text revisions due to its small sample size. With only 10 participants, the relative effects of peer feedback, particularly those aspects of peer feedback related to students' language proficiency levels need to be cautiously interpreted.

A recent study, conducted by Dressler, Chu, Crossman and Hilman (2019), compared the quantity and quality of uptake for peer feedback and teacher feedback for graduate students. In their study, a cohort of nine participants engaged in asynchronous group peer feedback activities and received teacher feedback when they finished their first

drafts of two writing tasks. Their findings were that the uptake rate for teacher feedback was higher than that for peer feedback (89.26% vs. 82.95%) and the quality of uptake for teacher feedback was higher than that for peer feedback (85.87% vs. 73.5%), but neither difference was statically significant. In addition, the participants were found to incorporate surface-level feedback more frequently and meaning-level feedback less frequently, but the difference was not statistically significant. With such findings, the researchers concluded that teacher feedback and peer feedback may be equally valuable for revisions and meanwhile cautioned against some factors that may have influenced the results, including the order of implementing different feedback forms, insufficient peer feedback training, the participants' specializations and their tendency to focus on surface problems, and the small sample size.

#### 2.3.1.2 *L2 studies on effects of peer feedback on text revisions*

Within the field of L2 peer feedback studies, effects of peer feedback on text revisions have long been a popular topic and researchers have employed a wide variety of measures in investigating its effects. Existing L2 studies on this topic can be categorized into two strands, depending on whether the studies included a control/comparison group.

The majority of early L2 studies on the impact of peer feedback on text revisions involved one cohort of participants who received peer feedback only or received peer feedback at one stage and other forms of feedback at other stages (e.g., Connor & Asenavage, 1994; Mendonça & Johnson, 1994; Min, 2006; Paulus, 1999; Tsui & Ng, 2000; Villamil & de Guerrero, 1998). In these studies, the two measures indicating utilization of peer feedback, the percentage of revisions triggered by peer feedback and incorporation rate of peer feedback, are most frequently employed. The findings, however, were disparate. For instance, Connor and Asenavage (1994) found that only

5% of the total revisions, made by eight participants, resulted from peer feedback, whereas Mendonça and Johnson (1994) claimed that 53% of the revisions were due to peer feedback. As regards incorporation rate of peer feedback, Tsui and Ng (2000) found that 21 of their 27 participants incorporated more than 50% of teacher comments but less than 50% of peer comments; Villamil and de Guerrero (1998) found their learners utilized 82% of peer feedback points by either incorporating them directly or making changes triggered by them in their revised drafts. The disparate findings from the above-mentioned studies, however, were not surprising when consideration was given to the presence of teacher feedback within the same writing task. When only peer feedback was provided in the writing task, the incorporation rate of peer feedback, or the percentage of revisions triggered by peer feedback was relatively high and satisfactory, as reported in Mendonça and Johnson (1994) and Villamil and de Guerrero (1998). When learners received peer feedback and teacher feedback within one writing task, they paid more attention to teacher feedback than to peer feedback (e.g., Connor & Asenavage, 1994; Paulus, 1999). In addition, the low percentage of revisions triggered by peer feedback and low incorporation rate of peer feedback may be attributed to the mode of peer feedback being oral in some studies (Connor & Asenavage, 1994; Mendonça & Johnson, 1994). When L2 students discussed each other's compositions on the spot, the multiple tasks of reading compositions, giving and receiving feedback may have prevented them from making proper notes of some feedback points. Such feedback points, compared with written teacher feedback points, were more likely to be missed in revision.

Another measure, the percentage of meaning changes in revisions triggered by peer feedback, was also frequently employed. Drawing upon Faigley and Witte's (1981) taxonomy of revision changes, researchers categorized revisions into surface changes

and meaning changes. Surface changes include formal changes and meaning-preserving changes which alter the surface structure without adding new information or deleting old information; meaning changes include microstructure and macrostructure changes which affect the information present in the text. Findings on this measure have been consistent and indicative of the positive effects of peer feedback on text revisions: Connor and Asenavage (1994) reported that 70% of revisions triggered by peer feedback were meaning changes and Paulus (1999) found meaning changes accounted for 63.3% of revisions triggered by peer feedback.

In later studies addressing the effects of peer feedback on text revisions, researchers modified existing measures and introduced new ones. Hu and Lam (2010) refined the measure “the incorporation rate of feedback” into “the incorporation rate of valid feedback” and included a new measure “the percentage of valid feedback received”. Zhao (2010) incorporated the percentage of feedback understood and found 83% of peer feedback points, in contrast with 58% of teacher feedback points, were incorporated with evidence of understanding. Although Zhao (2010) employed both the incorporation rate of feedback and the percentage of feedback points understood in her study, she claimed that the latter measure was more effective, with the argument that understanding is a prerequisite for using feedback. Villamil and de Guerrero (1998) reported that self-revisions accounted for 39% of the total revisions that their participants made.

Hu and Lam’s (2010) study, using revision measures (the proportion of valid peer feedback points received, proportion of valid peer feedback points taken up in revision) and writing scores to investigate effects of peer feedback on text revisions is a representative for early studies. They reported that around 75% of peer feedback points received were valid and that around 75% of valid peer feedback points were taken up in



revision. They also found their participants made significant gains in overall score, content and organization score, and language score between drafts. Moreover, analyses of correlations between gains in writing scores and feedback measures suggest that the improvements in the revised drafts may have been linked to peer feedback. This study, in comparison with others in this strand, has two merits. Firstly, it incorporated an overall score and two sub-scores, showing how different dimensions of writing were impacted by peer feedback. Secondly, employment of correlational analysis compensated, to some extent, the methodological issue of not including a comparison group.

Kamimura's (2006) study, conducted in a Japanese university, examined effects of peer feedback on text revisions for learners with different EFL proficiency levels. By comparing the two drafts of an argumentative writing task using two measures, overall score and word count, the researcher found that learners with low EFL proficiency level made greater gains in overall score than those with high EFL proficiency level, and only learners with low EFL proficiency level showed an increase in word count. While the findings for changes in overall scores are consistent with that reported by Lindgren (2005), suggesting the potential of peer feedback in improving the writing of learners with low language proficiency, there is a methodological issue in this study. Because the data were collected from the same writing task in which peer feedback training was also implemented, peer feedback training may have led to improvements that learners made between drafts.

The other strand of L2 studies that investigated effects of peer feedback on text revisions features a quasi-experimental design and involves a larger number of participants.

Typically, these studies included a comparison/control group and relied on statistical analysis in generalizing the effects of peer feedback.

In the Chinese EFL context, M. Yang, Badger, and Yu (2006) compared effects of peer feedback and teacher feedback on university students' text revisions. Two intact classes participated in this study, with one class engaging in paired peer feedback activities and the other class receiving teacher feedback for three writing tasks. The researchers collected textual and questionnaire data from all participants in the two classes, as well as video recording and interview data from a total of 12 case study participants (six from each class). By comparing the two drafts written by all participants in terms of overall score, the researchers found that students in the teacher feedback class improved more than their counterparts in the peer feedback class. For the analysis of case study data, the researchers used four revision measures: the incorporation rate of usable feedback (adapted from the incorporation rate of feedback), the percentage of meaning changes, the percentage of successful changes, and frequency of self-revisions. Of the four measures, the case study participants who had teacher feedback outperformed their counterparts from the peer feedback class in only one measure, incorporation rate of usable feedback. The case study participants who had peer feedback, however, had a higher percentage of meaning changes (27% vs 5%) and a higher percentage of successful revisions (98% vs 87%), and made more self-revisions (16 vs 5) than those from the teacher feedback class. As M. Yang et al. (2006) included a comparison group in their study, their finding was more convincing than those from the studies that implemented different forms of feedback to just one group of participants. Furthermore, the researchers, by utilizing multiple measures and employing mixed methods, showed the indispensable role that peer feedback played in revision.

Conducted in a Japanese university, Ruegg (2015b) examined students' uptake of peer feedback and teacher feedback. Based on a random collection of feedback forms and drafts from the 64 participants in a semester-long writing course, Ruegg (2015b) found that teacher feedback was more often incorporated in revision than peer feedback, but teacher feedback was more often misunderstood than peer feedback. In terms of revision quality, peer feedback led to successful revisions more frequently than teacher feedback. Regarding the differences in uptake of feedback, the researcher attributed them to two factors: feedback specificity and power relations. The teacher feedback group participants benefited from receiving coded marks and general comments which contained more information about revision; the trust they had in their teachers also led to a high incorporation rate of teacher feedback. The power distance between teachers and students prevented the participants from asking questions and therefore led to some misunderstanding of teacher feedback and unsuccessful revision attempts. For the peer feedback group participants, they had a low incorporation rate of feedback because they received some unspecific feedback, but they had a higher understanding rate of received feedback and a higher percentage of successful revisions because the equal status among peers allowed them to critically evaluate the received feedback points and selectively use them in revision. Compared with previous studies, the student-led feedback form in Ruegg (2015b) was an innovation. The feedback form included four questions from student writers and an additional question asking peer/teacher reader to provide one piece of constructive feedback. When student writers received feedback points they needed, it was more likely they would pay attention to them when writing their revised drafts.

Two studies (Diab, 2010, 2011) examined effects of peer feedback and self-feedback on text revisions. Forty Lebanese university EFL students participated in this study, with

22 students in the peer feedback group and 18 students in the self-feedback group. The two groups received instruction on argumentative essays and the four grammar structures targeted in the study and had peer feedback or self-feedback training. In Diab's (2010) study, the first drafts and the revised drafts of two writing tasks were compared respectively. Students in the peer feedback group were found to reduce rule-based errors (subject/verb agreement and pronoun agreement errors) more than those in the self-feedback group, but neither group significantly reduced non-rule-based errors (inappropriate words and awkward sentence structures). Diab's (2011) second study investigated the impact of peer feedback and self-feedback on changes in the overall quality between drafts: Students in the peer feedback group made greater improvement in the overall quality than those in the self-feedback group. Additionally, Diab (2011) integrated quantitative and qualitative findings by analyzing feedback points collected from the 14 case study participants (seven from each group) and the corresponding changes in their revised drafts. The researcher found that students in the two groups similarly detected content and organization problems in compositions but differed in revision quantity and quality. Students in the peer feedback group made fewer revisions than those in the self-feedback group but had a higher percentage of revisions that improved text quality of their revised drafts.

Although early L2 peer feedback studies included only one group of students and relied on descriptive statistics in examining effects of peer feedback, later studies have included a comparison group and employed multiple instruments. Nevertheless, since improvements in revised drafts serve only as "tentative evidence of actual growth" (Villamil & de Guerrero, 1998, p. 503), there is a need to examine new writing to establish the efficacy of peer feedback.

### 2.3.2 Studies on effects of peer feedback on writing performance

To examine effects of peer feedback on writing performance, researchers generally adopted a quasi-experimental design, with a pre-test and a post-test, and several peer feedback practices in between as an intervention. In the following subsections, L1 and L2 studies examining the effects of peer feedback on writing performance will be reviewed respectively.

#### 2.3.2.1 *L1 studies on effects of peer feedback on writing performance*

L1 studies examining the effects of peer feedback on writing performance have been very scarce. Conducted in primary and secondary settings, these studies generally reported that peer feedback was more effective in helping students improve their writing ability.

Karegianes, Pascarella and Pflaum (1980) implemented a highly-structured peer editing program in which low-achieving tenth grade students “were carefully taught procedures for editing each other’s work” (p. 206). They found that students in peer editing group had significantly higher writing proficiency than those in teacher editing group after 10 weeks.

Prater and Bermudez (1993), however, reported some different results. With the intention of providing low-achieving elementary school students with opportunities to develop language and social skills, Prater and Bermudez (1993) organized students in the experimental group into heterogeneous peer response groups and incorporated peer response activities in multiple stages of each writing task, including prewriting, drafting and revising. After three writing tasks with peer response activities, students in the peer response group made similar gains in overall scores as those in the teacher feedback

group, but they made greater gains in fluency than those in the teacher feedback group, as measured by the number of words and clauses. Findings from this study need to be interpreted with caution because of its short duration: “One month is not a sufficient length of time to produce significant differences in overall quality of writing” (Prater & Bermudez, 1993, p. 108). Clearly, a study with a longer intervention period might reveal the advantage of peer feedback over teacher feedback. Meanwhile, there are other points to note for this study. Firstly, although the study was conducted in L1 primary school classrooms, all participants used to be in ESL or bilingual classroom at one time. Secondly, the genres for writing tasks during the intervention may have differed from those for the pre- and post-tests, because students decided for themselves what topics to write on. Thirdly, neither the pre-test or the post-test was conducted during class hours, with time constraints and specific requirements.

Conducted in a Belgian secondary school (Dutch as the first language), Gielen et al. (2010) examined the effectiveness of two extended forms of peer feedback in improving students’ overall writing performance. Their study included three writing assignments and involved three peer feedback groups and one teacher feedback group. The conventional peer feedback group only used a peer feedback form listing several criteria for evaluating their peers’ essays, while the question-peer feedback and peer feedback-reply groups had an additional form, eliciting the writer’s expectation for peer feedback and the writer’s response to peer feedback respectively. The teacher feedback group received collective feedback, a practice in which the teacher addressed some common issues to the whole class based on his/her evaluation of some collected essays. Like Ruegg (2015a), this study only included written feedback mode for the four groups. By comparing the students’ overall scores in the pre-test and post-test, the researchers found that the two extended forms of peer feedback were superior to traditional peer feedback

or collective feedback in helping students improve in writing. Traditional peer feedback and collective feedback had similar effects on improving students' writing performance.

Crinon and Marin (2010) and Crinon (2012) compared the respective effects of giving feedback and receiving feedback on development of writing skills for Grade 4 and Grade 5 students with different L1 (French) proficiency level. Crinon and Marin (2010) implemented four explanatory writing tasks with giving/receiving feedback activities within one academic year and found students in the feedback giving group wrote longer texts, with a larger amount of relevant information and a smaller amount of irrelevant information than those in the feedback receiving group. Similarly, Crinon (2012) focused on narrative writing tasks and reported that students in the feedback giving group wrote longer texts, with a larger amount of relevant information and coherent descriptive information, than those in the feedback receiving group. Apart from the statistical analysis, Crinon (2012) included a case study to showcase the development of texts for each group. After the four narrative writing tasks, the case study participant from the feedback giving group produced texts in compliance with success criteria of narrative writing, while the case study participant from the feedback receiving group did not move beyond his initial level. For the better performance in writing for the feedback giving group, Crinon and Marin (2010) explained that providing feedback allowed students to build knowledge about the practiced genre and to apply suggestions they gave others to their own writing. For the benefits of giving feedback, Cho and MacArthur (2011) explained the contribution of its specific processes: detecting and diagnosing problems and recommending solutions helped students practice revision and writing skills and led to improved understanding and giving explanations provided students with opportunities of clarifying their understanding.

There are, however, two issues to note for Crinon and Marin (2010) and Crinon (2012). The findings related to language proficiency were scarce and vague. Crinon and Marin (2010) explained that the weaker students in the feedback giving group made as much progress as the stronger students, while the weaker students in the feedback receiving group had the most difficulty when writing on the last topic; Crinon (2012) mentioned that “In the group of participants giving advice, good students in particular show signs of improvement on several criteria” (p. 144), without exemplifying the criteria. Also, the two studies are low in ecological validity, because they did not reflect reciprocal peer feedback activities as the norm in writing classrooms (Cho & MacArthur, 2011). In writing classrooms, there is the need to implement reciprocal peer feedback activities to help students develop their writing skills (Cho & MacArthur; 2011; Crinon, 2012; Lundstrom & Baker, 2009).

#### *2.3.2.2 L2 studies on effects of peer feedback on writing performance*

In the recent ten years, there have been a growing number of L2 studies investigating the effects of peer feedback on writing performance, with conflicting results documented (e.g., Gai & Zhou, 2013; Ruegg, 2015a; L. Yang, Yang, & Zhou, 2013; X. Zhang & McEneaney, 2020; Y. Zhou, 2013). Conducted in the Japanese university context, Ruegg (2015a) compared the effects of peer feedback and teacher feedback on writing performance after a one-year intervention with 16 feedback iterations. For each draft of a writing task, the participants raised four questions for feedback and received written comments from their peer/teacher. The findings were that participants in the peer feedback group made similar gains in content, organization, vocabulary, and overall scores as those in the teacher feedback group, but smaller gains in grammar scores than those in the teacher feedback group. In other words, except for grammar, peer feedback was as effective as teacher feedback in helping students improve in writing over time.



With ESL students in writing classes at a US university, Lundstrom and Baker (2009) investigated the effects of giving feedback and of receiving feedback on development in writing abilities and considered the effects of L2 proficiency levels (beginning and intermediate). They assigned students at each proficiency level to the feedback giving group and the feedback receiving group. A comparison of their essays written at the beginning and end of the semester indicates that for the beginning students, the feedback giving group made greater gains in overall, development (appropriate use of examples and support), organization (effectiveness of the thesis statement and unity of ideas), and cohesion scores than the feedback receiving group. For the intermediate students, the feedback giving group made greater gains in overall, development and organization scores than the feedback receiving group. This study corroborates as well as advances Crinon and Marin (2010) and Crinon (2012) in that the results confirmed the value of giving feedback for L2 learners and revealed the effects of language proficiency levels. This study, however, had two ecological issues to consider. Similar to Crinon and Marin (2010) and Crinon (2012), the feedback giving/feedback receiving task, as implemented in this study, diverged from the norm in the writing classrooms. Also, this study used essays written by previous students rather than authentic writing tasks. Such a practice may have taken up time that students could use to practice their writing.

Studies conducted in the Chinese EFL context have revealed how different peer feedback forms impacted learners' writing performance. With undergraduate students as their participants, both Gai and Zhou (2013) and X. Zhang & McEneaney (2020) found students in the peer feedback group made greater improvements in overall writing quality than those in the comparison group. The two studies, however, differed in their interventions. Gai and Zhou (2013) implemented heterogenous-group peer feedback activities for a semester and provided students with peer feedback training prior to the

intervention. X. Zhang & McEneaney (2020) employed paired peer feedback activities for 15 weeks (approximately one semester) and incorporated training prior to the intervention and well-designed peer feedback activities within each writing task (peer feedback and author response instruction, teacher modelling and student practice, teacher-student conference and using author response forms). Similarly, Y. Zhou (2013) compared the effects of teacher feedback, traditional peer feedback and online peer feedback on undergraduates' writing performance. The results were that teacher feedback, peer feedback, and online peer feedback had similar effects on improving Chinese EFL learners' overall writing ability, but their contributions to the four dimensions of writing varied: Teacher feedback was most effective for improvements in organization but least effective for improvements in vocabulary; online peer feedback was most effective for improvements in content but least effective for improvements in grammar and mechanics. Another study, L. Yang, Yang, and Zhang (2013), reported that students receiving a combination of teacher and peer feedback made the greatest gains in the overall score and sub-scores (content, organization, grammar, vocabulary and mechanics) from pre-test to post-test, while students receiving peer feedback made the smallest gains in each of these scores.

The above-mentioned studies all featured a long intervention period, with Ruegg (2015a) being most representative for spanning one year and including 16 feedback iterations. The mixed findings, however, may be attributed to various factors. First, feedback quantity may have some influence on the results. In L. Yang, Yang, and Zhang (2013), the participants in the combined peer-teacher feedback group, receiving feedback from two sources, may have received most feedback points. The teacher feedback group participants may have received more feedback points than those in the peer feedback group because of they had a collective feedback session in which the teacher

summarized common problems detected from the group participants' drafts. The peer group participants, however, only discussed problems in their own draft and their peer's drafts after receiving written feedback from their peers. Similarly, peer feedback groups with four members in Gai & Zhou (2013) may have provided students with more feedback points than those in the teacher feedback group, contributing to their better performance in writing after the intervention. Secondly, feedback focus and quality may help explain the different effects of peer feedback and teacher feedback. Ruegg (2015a) found students in the teacher feedback group received a high percentage of meaning-level feedback and they made greater improvements in grammar of writing. For the little improvement made by the peer feedback group, Ruegg (2015a) explained that some surface-level grammar feedback being inaccurate was a possible reason. Thirdly, feedback mode may have been another factor contributing to the different effects of peer feedback on writing performance. Theoretically, one advantage for peer feedback originates from collaborative learning opportunities in peer interactions and negotiations. When Ruegg (2015a) implemented written peer feedback in her study, the benefit of peer interactions and negotiations was missing. Consequently, the effects of peer feedback on overall and some sub-dimensions of writing performance may have been compromised. Finally, whether learners were provided with sustained support throughout the intervention may have some influence on the results. Compared with the one-shot training prior to the intervention, common in most peer feedback studies, the on-going support offered by X. Zhang & McEneaney (2020) may have contributed to the better effects of peer feedback on writing performance.

Another study, Diab (2016), conducted in the Lebanese EFL context, compared the effectiveness of three types of feedback (teacher feedback, peer feedback, and self-feedback) in reducing the two types of errors, pronoun agreement errors and lexical

errors. This study compared the three group participants' compositions in the pre-test (Week 1, before the intervention), in the post-test (Week 8, immediately after the intervention), and in the delayed post-test (Week 16, eight weeks after the intervention). The research findings showed that the three feedback forms were equally effective in reducing pronoun agreement errors in the long-term. Participants in each group increased in pronoun agreement errors in the post-test, but they reduced in such errors in the delayed post-test; no significant difference existed between any two groups in the post-test and delayed post-test. For lexical errors, only the self-feedback group significantly reduced in its percentage over time. Although the percentage of lexical errors for the peer feedback group increased in the post-test and then reduced in the delayed post, neither change was statistically significant. To account for the participants' disparate performances in the post-test and in the delayed post-test, Diab (2016) raised several methodological issues to consider. Firstly, the pre-test, post-test, and delayed post-test differed in difficulty level and in format, with the post-test writing task being the most difficult one of the three tests and the delayed post-test being a graded writing task for which participants' increased motivation due to grading may have yielded improved performance in the two types of errors. Secondly, the post-intervention task involved "extensive self-monitoring of language production" (Diab, 2016, p. 62), which may have a training effect on the participants' ability to detect linguistic errors in their writing, and therefore leading to their improved performance in the delayed post-test. Although the above-mentioned issues may have impacted validity of the research findings, the pre-, post-, and delayed post-tests design in Diab's (2016) study has informed the present study by suggesting a new time dimension for peer feedback research. As more and more researchers have employed such a design in teacher feedback studies (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010; Rastgou,

Storch, & Knoch, 2020; Sheen, Wright, & Moldawa, 2009; Shintani & Ellis, 2013), a delayed post-test needs to be included in peer feedback studies.

In a recent study, Shang (2019) incorporated some complexity, accuracy and fluency (CAF indexes) when investigating the effects of online peer feedback on writing performance. Conducted in a 18-week writing course, Shang (2019) alternately implemented online peer feedback and automated corrective feedback activities with the 47 first-year English majors at a Taiwan university. The results were that online peer was more effective than automated corrective feedback in helping students improve in fluency, accuracy, and lexical complexity, as seen from changes in the number of sentences, errors and word types respectively. As the first attempt to incorporate CAF indexes in peer feedback study, Shang (2019) provided empirical evidence for the impact of peer feedback on different sub-dimensions of language use in writing. There are, however, two research design problems in this study. Firstly, since this study only involved four writing tasks (Tasks 1 and 3 with peer feedback activities and Tasks 2 and 4 with automated corrective feedback activities), the participants may have had limited experience with either feedback form. Secondly, some CAF measures may have been inappropriately used. Of the measures used in this study, four measures, the number of errors, word types, word tokens, and the type-token ratio, were very sensitive to text length (Ellis & Barkhuizen, 2005). The findings, therefore, may not reliably reflect changes in the participants' linguistic performance.

### **2.3.3 Training in peer feedback studies**

Although some researchers included minimal training in L1 peer feedback studies or made no explicit mention of training in their articles (e.g., Cho & MacArthur, 2010; Graner, 1987), there has been a shared opinion that training is an essential component

in peer feedback studies (See Chang, 2016; Yu & Lee, 2016). To facilitate implementation of peer feedback training, researchers offered practical suggestions on training activities and procedures. For instance, McGroarty and Zhu (1997) introduced teacher-student conferences in peer feedback training; Min (2005) incorporated the four-step procedure for constructing feedback, “clarifying the writer’s intention, identifying problems, explaining the nature of problems, and making suggestions by giving specific examples”, in training (p. 293); Hu (2005) presented a variety of practical activities for peer feedback training; and Bui and Kong (2019) incorporated metacognition in peer feedback training and designed training activities accordingly. Most studies, have revealed, empirically, various benefits of peer feedback training. Peer feedback training helped students improve in feedback quantity and quality, incorporate a higher percentage of peer feedback in revision, improve in writing, as well as take a collaborative stance in peer feedback activities (McGroarty & Zhu, 1997; Min, 2005, 2006, 2008; Rahimi, 2013; Stanley, 1992).

While the majority of existing studies implemented peer feedback training before students embarked on reviewing activities, some researchers noted the limitation of “one-off” pre-intervention training (Lockhart & Ng, 1995, p. 648). Since such training typically involved sample feedback and sample compositions, it “may not be able to offer immediate support to reviewers in actual feedback tasks” (Lee, 2015, p. 3). Consequently, researchers advocated other forms of training activities during the intervention, such as teacher-student conferences (Min, 2008) and a teacher’s ongoing guidance about feedback and author response strategies (X. Zhang & McEneaney, 2020; Q. Zhu & Carless, 2018).

### 2.3.4 Summary of findings on effects of peer feedback

L1 and L2 studies have presented mixed results regarding effects of peer feedback on students' text revisions. L1 studies examining the effects of peer feedback on text revisions have been conducted in different school settings, with a comparison group receiving regular teacher feedback or self-feedback. In comparison, the majority of L2 peer feedback studies examining the effects of peer feedback on text revisions have been in university settings, with some early studies including no control/comparison group. They employed multiple measures and reported divergent results. Seen from the measures indicating utilization of feedback in revision, peer feedback was not as effective as teacher feedback for revision (Connor & Asenavage, 1994; Lam, 2013; Ruegg, 2015b; Tsui & Ng, 2000; Zhao, 2010). Peer feedback, however, played a positive role in text revisions since it led to a high percentage of meaning changes (Paulus, 1999; M. Yang et al., 2006), and triggered a substantial number of self-revisions (Villamil & de Guerrero, 1998). Students were also found to have less difficulty in understanding peer feedback than teacher feedback (Ruegg, 2015b; Zhao, 2010). Later L2 studies investigating the effects of peer feedback on text revisions have included a comparison group. The limited findings, however, suggest that peer feedback was less effective than teacher feedback, but more effective than self-feedback, in helping L2 students improve in overall score between drafts (Diab, 2011; Ruegg, 2015b; M. Yang et al., 2006).

Compared with revision studies, there have been few investigations into the effects of peer feedback on writing performance. In L1 and L2 studies, researchers compared the effects of peer feedback with other feedback forms as well as different forms of peer feedback (e.g., Gai & Zhou, 2013; Gielen et al., 2010; Karegianes et al., 1980; Ruegg, 2015a; X. Zhang & McEneaney, 2020; Y. Zhou, 2013). Using writing scores and

linguistic indexes, these studies have revealed how peer feedback impacted different dimensions of writing performance, yet with divergent results. The findings from L1 studies generally suggest that peer feedback was more effective than teacher feedback or self-feedback in helping students improve in their overall writing ability as well as in their abilities to develop content and produce a longer text (Karegianes et al., 1980; Prater & Bermudez, 1993). Studies conducted in L2 contexts, however, implemented different forms and peer feedback and reported different results.

This review of the literature, however, uncovered some methodological problems. The frequency that learners are exposed to peer feedback tasks needs to be increased. Although most of existing studies, particularly those examining the effects of peer feedback on writing performance, featured a long intervention period, they differed in feedback frequency. Based upon analysis of 24 feedback studies, either self-identified or identified by a second source as longitudinal or long-term studies, Ruegg (2020) suggested that studies examining effects of feedback be conducted over an extended period (at least one semester) and involve more feedback iterations (four feedback iterations and above).

Another issue that arises is that some measures in existing studies are used inappropriately. Some measures of feedback and revision, such as the incorporation rate of feedback and the percentage of revisions triggered by feedback, showing quantity and quality of specific revisions, may not reflect text qualities. In Diab (2016), person agreement error is a much narrower concept than lexical error; grammatical error, an equivalent concept to lexical error may be more appropriate. Some CAF indexes (e.g., the number of errors and type-token ratio) used by Shang (2019) are problematic in that they were very sensitive to text length. Yu and Lee (2016) suggested a combined use of



writing scores and CAF indexes, in that CAF indexes were objective quantitative measures, capable of showing subtle changes in language development (Housen, Kuiken, & Vedder, 2012; Wolfe-Quintero, Inagaki, & Kim, 1998). When used to evaluate quality of a composition, they can show whether or not the participants made improvements in language use, and more importantly, the dimensions in which they made the improvements.

A further issue is that little attention has been paid to the long-term impact of peer feedback. Of existing peer feedback studies, only Diab's (2016) study included a delayed post-test. As Yu and Lee (2016) have identified the need to investigate "the long-term impact of peer feedback on students' L2 writing development" (p. 483) and an increasing number of teacher feedback studies have employed the pre-, post-, and delayed post-tests design, a delayed post-test should be included in peer feedback studies.

Finally, training is essential for peer feedback studies. Based on literature on implementing peer feedback training and empirical findings about effects of peer feedback training, research, which includes peer feedback training and on-going support to achieve the optimal effect of peer feedback, is needed.

## **2.4 Processing of peer feedback in revision**

Within the field of peer feedback studies, there is a lack of research on students' processing of peer feedback in revision. In this section, major revision models are described, followed by a review of empirical studies on students' processing of peer feedback and teacher feedback.

## 2.4.1 Major models of revision

The establishment of revision as a legitimate research area was shaped by Flower and Hayes' (1981) cognitive process writing model which views revision as an essential process. Since the 1980s, several revision models have been proposed, including Scardamalia and Bereiter's (1983) compare/diagnose/operate (CDO) revision model, Hayes, Flower, Schriver, Stratman, and Carey's (1987) cognitive process model of revision, Hayes' (1996) new cognitive process model of revision, and Butterfield, Hacker and Albertson's (1996) cognitive-metacognitive model of revision. In the following subsections, the key elements of each model are identified and discussed.

### 2.4.1.1 *Scardamalia and Bereiter's (1983) CDO revision model*

With the purpose of helping children revise, Scardamalia and Bereiter (1983) proposed a revision model which described the three steps in revision: comparing, diagnosing, and operating. According to this model, writers keep two types of representations in mind when writing, the intended meaning to be conveyed through the text and the actual text produced. First, writers compare and evaluate the two representations, searching for discrepancies between them. When a mismatch is detected, they diagnose the nature of the problem and consider possible solutions. They, then select a strategy (e.g., deleting, rewriting) to solve the problem and generate a new text.

The value of the CDO revision model lies in pinpointing four essential processes in revision: comparing, diagnosing, selecting a strategy, and generating a new text. This model, which leads writers through the four essential processes and avoids short-circuiting during revision, is criticized, however, for oversimplification. Firstly, it focuses narrowly on the writer's cognitive processes, giving no consideration to other factors in revision. Secondly, its underlying assumption that revision occurs only when

writers detect problems in their texts is problematic. In practice, there are many instances in which writers revise because they discover a better way to express their ideas.

#### 2.4.1.2 *Hayes et al.'s (1987) cognitive process model of revision*

Hayes et al.'s (1987) cognitive process model of revision is an elaboration of the reviewing process in Flower and Hayes' (1981) cognitive process writing model (see Figure 2.3). This model, unlike the prescriptive CDO revision model, is based on L1 writers' think-aloud protocols for revision.

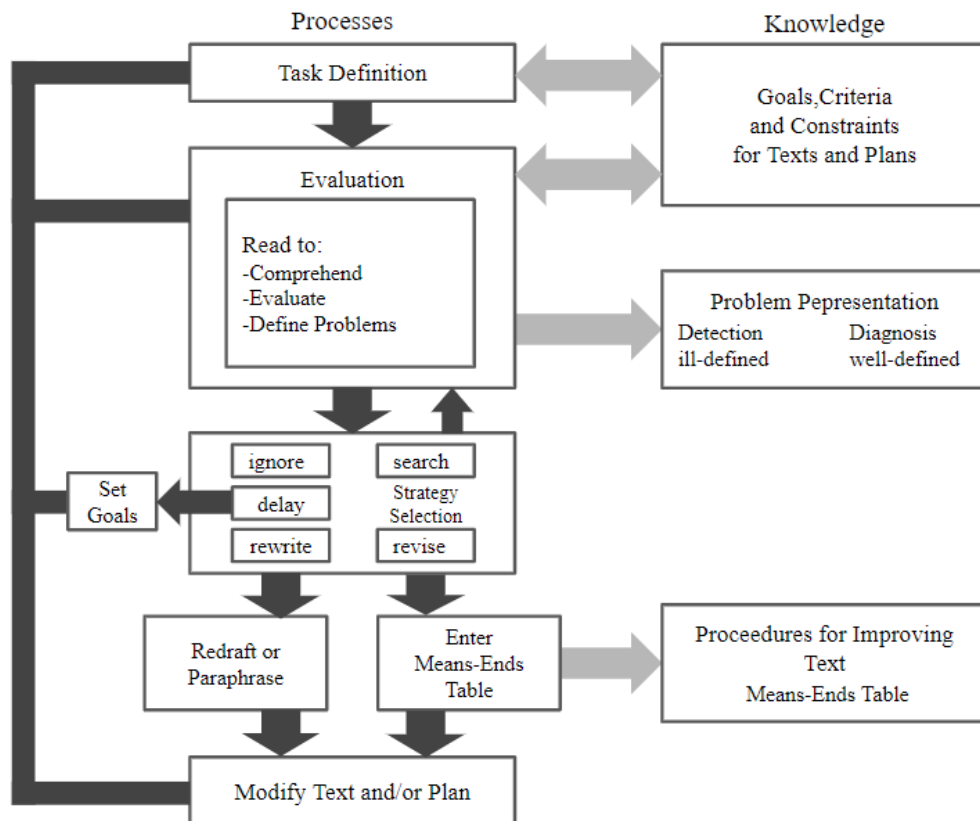


Figure 2.3 The Hayes et al.'s (1987) cognitive process model of revision

The model has two major components: Processes writers undergo and categories of knowledge that influence these processes. The categories of knowledge include declarative knowledge about the features of a problem (goals, criteria and constraints for texts and plans) and procedural knowledge regarding how to address the problem (problem representation and procedures for improving text). Task definition is a process in which a writer plans what to revise and how to carry out the revision. Within this model, task definition is critical in that it “guides the whole revision process and determines the sequence of processes” (Dost, 2019, p. 10).

The evaluation process starts after the writer sets the goals and plans during the task definition stage. Within the evaluation process, three types of reading are distinguished. Reading to comprehend is least demanding since the writer simply reads a text to understand its information; reading to evaluate involves intentional assessment against goals or criteria and detection of text problems, while reading to define problems goes beyond assessment to diagnosis. The primary outcome of evaluation is problem presentation, as a continuum ranging from being ill- to well-defined. Based upon the problem representation, the writer chooses from the five strategies (ignoring the problem, delaying action, searching for better representation of the problem by further evaluation, rewriting, and revising) to deal with the problem.

In summary, this model, in comparison with Scardamalia and Bereiter’s (1983) CDO revision model, demonstrates the complexity of revision, “with hierarchically organized sub-processes, that serially appear, or that are subordinated to other processes” (Alamargot & Chanquoy, 2001, p. 108). It also highlights the interaction between the two processes (task definition and evaluation) and the knowledge component. Although

this model does not consider cognitive resources or contextual factors in revision, it serves as the central framework for subsequent revision models.

#### 2.4.1.3 *Hayes' (1996) cognitive process model of revision*

In 1996, Hayes developed an updated model of writing with a submodel for revision, Hayes' (1996) cognitive process model of revision (see Figure 2.4). This model comprises three parts: the control structure, fundamental processes, and resources. In the control structure, the task schema is defined as “a package of knowledge, acquired through practice, that is useful for performing the task and is retrieved as a unit when cues indicating the relevance of the schema are perceived” (Hayes, 1996, p. 16). Specifically, it includes “a goal, several revising activities, attentional goals, revising criteria, and strategies” (Chanquoy, 2009, p. 84). Fundamental processes have three components: text processing with critical reading, reflection with problem-solving and decision making, and text production. Finally, resources include working memory and long-term memory.

As an update for Hayes et al.'s (1987) revision model, this model features several developments. Firstly, task definition in the previous model is replaced by control structure, to emphasize that control structure greatly impacts revision performance (Chanquoy, 2009). Secondly, the emphasis is on reading and comprehension, as can be seen from the use of critical reading in this model. Critical reading, which differs from reading to comprehend, emphasizes the goal of identifying problems in the text, especially those concerning overall structure, coherence, and meaning of the text (Chanquoy, 2009). Thirdly, this model incorporates working memory and long-term memory, demonstrates the interaction between them, and distinguishes their roles in revision. The limited resources in working memory are exploited both for storing

information and for performing the fundamental processes which are not automated. In the long-term memory, the knowledge, including general topic knowledge and specific knowledge about writing, is stored. Such knowledge can be activated and utilized in the fundamental processes or be retrieved into working memory to participate in the revising process.

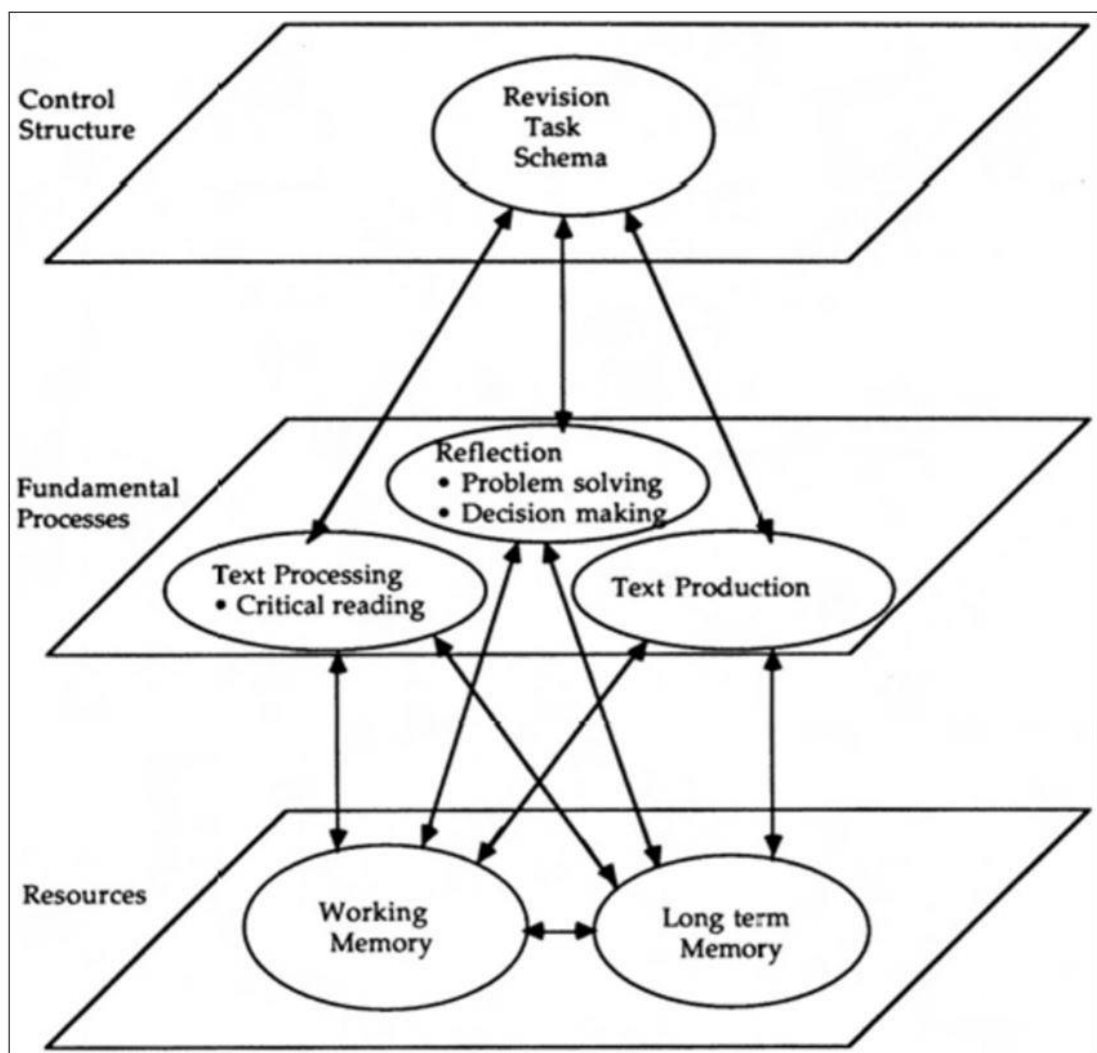


Figure 2.4 The Hayes' (1996) cognitive process model of revision

#### 2.4.1.4 Butterfield et al.'s (1996) cognitive-metacognitive model of revision

Butterfield et al.'s (1996) cognitive-metacognitive model of revision comprises two major parts: environment and cognitive/metacognitive system (see Figure 2.5). The environment, external to the reviser, includes the rhetorical problem (the topic, the audience and the importance of the text to be revised) and the actual text being revised (the format, genre, lexical units, syntactic units, propositions, and gist of the already written text).

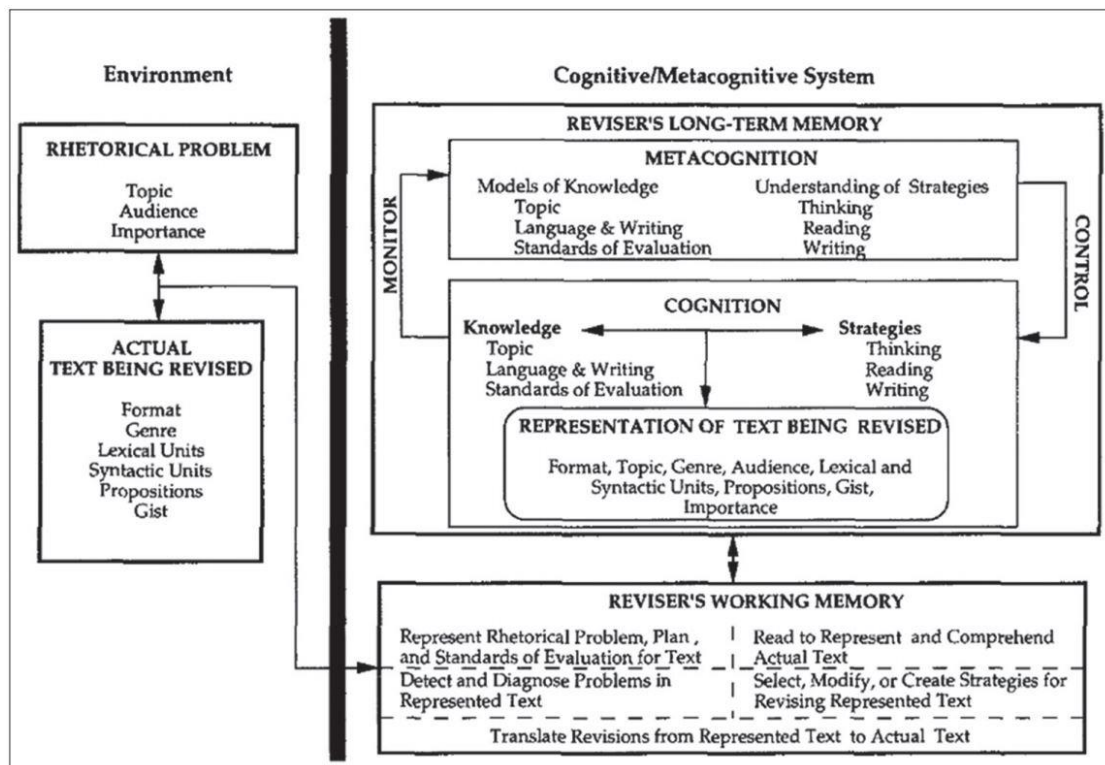


Figure 2.5 Butterfield et al.'s (1996) cognitive-metacognitive model of revision

The cognitive/metacognitive system has two subsystems: working memory and long-term memory. Butterfield et al. (1996) posited that the five fundamental revising

processes occurred in working memory, which include (1) representing the rhetorical problem, planning, and deciding on standards of evaluation; (2) reading to represent the text; (3) detection and diagnosis; (4) strategy selection for revising text; (5) translation from mental revision to actual text. Long-term memory comprises cognition and metacognition. The cognition section contains knowledge (topic, language and writing, and standards of evaluation), strategies (thinking, reading, and writing), and a representation of the text being revised. The metacognition section contains the same knowledge and strategies but in more synthesized forms, which are described as models of knowledge and understanding of strategies. Within long-term memory, there are interactions between cognition and metacognition through monitor and control functions. Monitoring includes “strategies such as re-reading a difficult part of the text, looking back to prior text, predicting the text-to-be-written, and comparing eventual solutions” (Dost, 2019, p. 13). With the control function, metacognition determines “when, where, how, and why it is necessary to use, evaluate, and control cognitive strategies and cognitive knowledge” (Chanquoy, 2009, p. 85).

This model, which differs from early revision models, illustrates the importance and roles of long-term memory, working memory and metacognitive knowledge in revision. It also incorporates audience “as part of the rhetorical problem the writers should consider related to the environment and as part of representation of text being revised” (Dost, 2019, p. 13). Moreover, this model features “interactions between different parts and/or between the different processing levels” (Alamargot & Chanquoy, 2001, p. 111).

#### 2.4.1.5 *Summary and critique of revision models*

The four revision models presented above, have demonstrated an evolution in order,



firstly, to describe as precisely as possible the different processes involved during revision, secondly, to consider the great difficulty of this activity, necessitating a writer's and a reviser's careful reading of the text in order to detect some errors or problems, and thirdly, to introduce long-term memory, working memory, and a metacognitive system in order to dynamically explore the revision process. (Chanquoy, 2009, p. 86)

The evolution in revision models has two major merits. Firstly, the four models are consistent in revealing fundamental processes of revision, though names for the processes vary. Secondly, while only Scardamalia and Bereiter's (1983) CDO revision model captures the fundamental processes, the subsequent ones keep incorporating cognitive factors and contextual factors in revision, providing a more comprehensive understanding about the revision process.

These models, however, have two critical weaknesses. First, feedback has not been considered in any of the four models, despite the indispensable role it plays in revision, especially for novice writers (Dost, 2019). Secondly, none of the models has deeply examined cognitive operations that writers use in each process. Of the four models, only Hayes' (1996) revision model includes three cognitive operations: critical reading in the process of text processing and problem solving and decision making in the process of reflection. These cognitive operations, however, may apply only to "expert writers...who are skilled in revision and use of cognitive sources"(Dost, 2019, p. 13). In other words, existing models may have failed to reflect how novice or EFL writers carry out each process in revision.

## 2.4.2 Empirical findings of L2 learners' processing of feedback in revision

There is little empirical research reporting students' processing of feedback in revision of drafts (see Storch & Wigglesworth, 2010). In this subsection, therefore, relevant findings from both peer feedback and teacher feedback studies that inform the present study are reviewed.

### 2.4.2.1 *L2 learners' processing of peer feedback in revision*

The only data that could be found on student' processing of peer feedback were from interviews with students or classroom observations. The case study participants in M. Yang et al. (2006) were reported to confirm the validity of peer feedback points by consulting reference books or their teachers before incorporating them in their writing. Villamil and de Guerrero (1996) suggested that some strategies utilized by their participants during peer interactions, such as employing symbols and external resources, using L1, and resorting to interlanguage (L2) knowledge, would also be used during revision. Yu, Zhang, Zheng, Yuan, and Zhang (2019) found their case study participants demonstrated superficial cognitive engagement with peer feedback in revision. Of the three case study participants, only one employed the meta-cognitive operation of monitoring, utilizing received peer feedback points to make unsolicited revisions, whereas the other two participants passively incorporated received peer feedback points in their revised drafts.

In a recent study, Fan and Xu (2020), based on data collected from semi-structured interviews with 21 Chinese non-English-major participants, reported a wide range of metacognitive and cognitive operations when processing received peer feedback points. These included monitoring, activating previous knowledge, reflecting (making judgements about the validity of a certain feedback point), using external resources, and

memorizing. They found the participants demonstrated extensive cognitive engagement with form-focused feedback, whereas they invested little effort in processing content-focused feedback. The researchers attributed the low engagement with content-focused feedback to “individual factors, such as low writing proficiency, priority to forms over content, and their past learning experience [their tendency to avoid the heavy workload involved in the processing content-focused feedback]” (Fan & Xu, 2020, p. 10). Their extensive engagement with form-focused feedback was explained by their instrumental motivation to pass CET-4, the most influential language proficiency test for non-English majors in China. Although Fan and Xu’s (2020) study identified how Chinese EFL learners processed peer feedback points and accounted for their different levels of cognitive engagement, two issues in research design need to be considered. Firstly, the use of semi-structured interviews might have prevented the researchers from fully capturing some cognitive operations because of the time lapse between the revision and interview. Secondly, in this study, no instruction was provided for how to use feedback in revision. Fan and Xu (2020) described there being a single focus on “strategies of evaluating a composition” in peer feedback training (p. 3), which might have influenced the participants’ engagement with peer feedback.

#### ***2.4.2.2 L2 learners’ processing of teacher feedback in revision***

A few teacher feedback studies have used verbal protocols to investigate L2 learners’ processing of feedback for revision of drafts. Qi and Lapkin (2001) used a case study to investigate how two L2 learners processed direct feedback, in the form of reformulation, received from a native speaker. The researchers analyzed the protocols and identified two cognitive operations: noticing (through reading the reformulated texts) and reasoning. It was observed that after the participants had noticed a language-related problem in their drafts as a result of receiving feedback, and had provided reasons for

their acceptance of feedback, they made changes in their revised drafts that led to improvement in text quality. Conversely, when they noticed a problem but failed to articulate a reason, they only occasionally made changes leading to improvement in text quality. This study has made a significant contribution because it identified that the cognitive operation of reasoning made a critical difference to the impact of feedback on text revisions. However, because the study focused narrowly on direct feedback, the findings may not generalize to other forms of feedback.

Storch and Wigglesworth (2010) collected feedback processing data from four paired discussions in which the eight students discussed written corrective feedback, in the forms of editing codes and formulations, that they received on their jointly-produced texts. They reported that the participants processed written corrective feedback in two ways, depending on feedback forms. In response to editing codes, they engaged extensively by identifying the nature of errors, activating their knowledge to offer suggestions and counter suggestions, and finally supplying the correct form whereas, in response to reformulations, they exhibited limited engagement by reading feedback and expressing agreement. Storch and Wigglesworth (2010), like Qi and Lapkin (2001), found extensive engagement with feedback contributed to uptake whereas limited engagement with feedback sometimes led to no uptake or inaccuracies. In addition, they suggested other factors that may have mediated the impact of corrective feedback, including the nature of errors and learners' attitudes, beliefs, and goals.

Based on the think-aloud protocols from eight ESL students, Rajoo (2009) proposed a model of thought processes about teacher feedback, including three processes: response, justification and revision (see Figure 2.6). In the response process, students responded to feedback by reading feedback, rereading feedback, referring to feedback, reading

their essays, rereading their essays and referring to their essays; they moved recursively among these operations. The response process was followed by justification, an optional process in which students provided justification for the received feedback. The justification process led to the last process in which students made their revisions by incorporating feedback or made no revision by rejecting feedback. This model clearly presents the processes that writers went through when revising in the presence of teacher feedback. In addition, it suggests that teacher feedback played an essential role in revision, and students moved back and forth between feedback and their texts in the response stage and occasionally “found justification for particular feedback before they responded either positively or negatively” (Rajoo, 2009, p. 149).

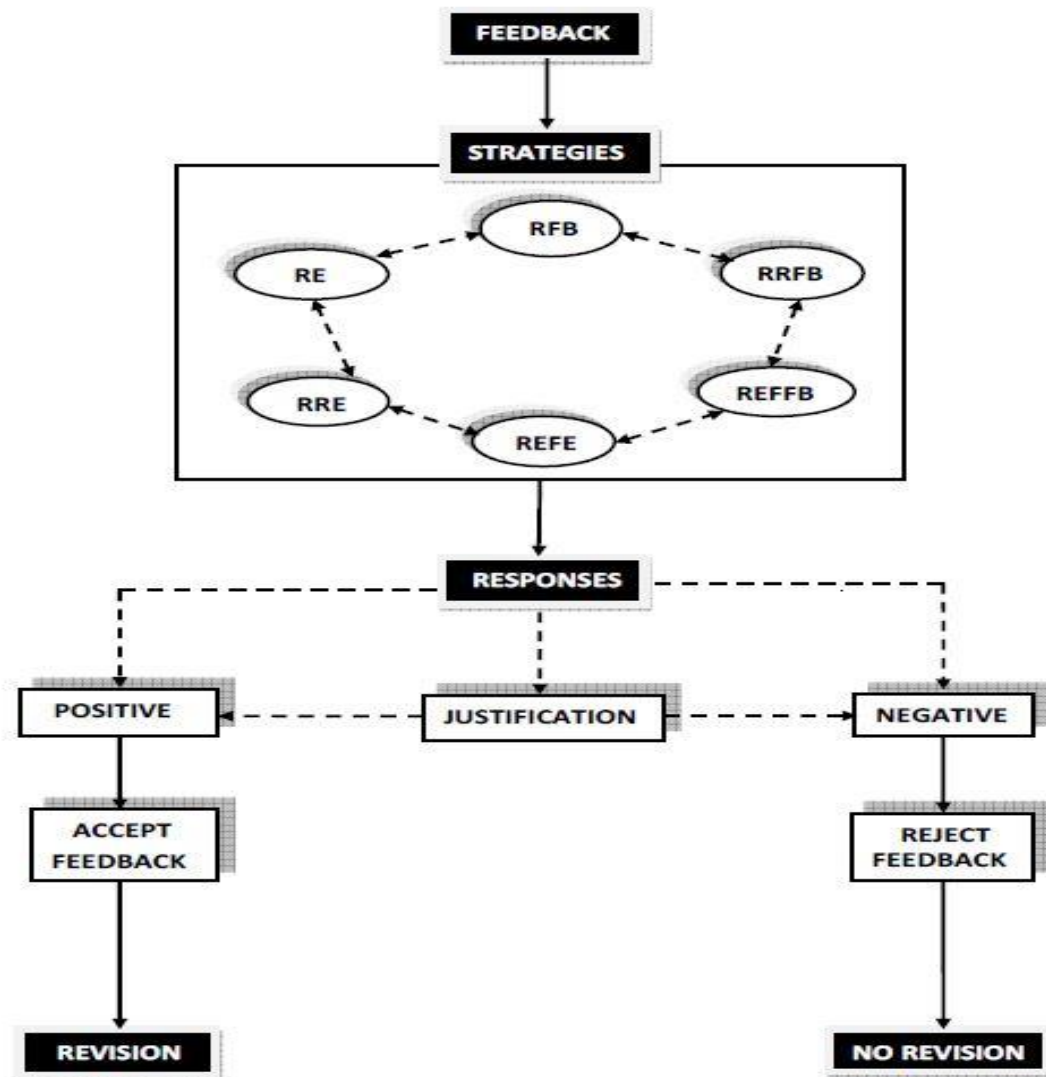


Figure 2.6 Rajoo’s (2009) model of thought processes on teacher feedback

*Note. RFB = Reads feedback (exact words); RE = Read parts of essay/sentences/section; RRE = Rereads or reviews parts of essay/sentences/section; REFE = Refer to parts of essay/sentences/section (not exact words); REFFB = Refer to feedback (not exact words); RRFB = Rereads or reviews parts of whole/parts of feedback.*

There are, however, several weaknesses in this model. Firstly, the cognitive operations in the response process are repetitive (“reading” and “rereading”) or unnecessarily fine-grained (“referring to” and “reading”), resulting in excessive strategies in the response

stage. Secondly, because only cognitive strategies for the response stage were included, this model gave the impression that other stages were less important. Thirdly, just as revision models showing no consideration of feedback would be incomplete (see 2.5.1.5), a processing model of teacher feedback without consideration of existing revision models would be similarly flawed. Since processing feedback and making revisions were inseparable, existing revision models need to be considered in research into processing of feedback in revision.

Two studies (Han, 2016; Yao Zheng & Yu, 2018) utilized retrospective verbal reports in investigating how Chinese EFL learners processed corrective feedback. Han's (2016) study, in comparison with Rajoo's (2009), yielded more cognitive and metacognitive operations. Cognitive operations included metalinguistic analysis, activating previous knowledge, predicting, making connections (by putting teacher feedback back to the text, by synthesizing different teacher feedback points, by linking teacher feedback to in-class instruction or other instruction materials), and memorizing teacher feedback and revisions, and metacognitive operations included managing attention, planning, and monitoring and evaluating. Students with low and high L2 proficiencies were found to use similar metacognitive and cognitive operations when processing teacher feedback in revision, but students with low L2 proficiency lacked the ability to orchestrate cognitive and metacognitive operations and to use internal and external resources to understand teacher feedback.

Yao Zheng and Yu (2018), however, found their twelve Chinese students with low English proficiency used only a few cognitive operations and virtually no metacognitive operation in processing corrective feedback. When revising, they simply read through their texts in conjunction with teacher feedback and made modifications when the

correct form was provided in the feedback. Such findings echo those in Yu et al. (2019) and could be attributed to two factors in research design. Firstly, the researchers asked their participants to self-report and record their response to teacher feedback without the presence of the researcher or instructor. Without training in verbal protocols, it is understandable that students with lower L2 proficiency may have reported superficially about their mental processes. Secondly, the researchers separated the concurring tasks of responding to teacher feedback and writing a revised draft into two stages. Thus, as the participants might not have seen any direct relevance of feedback to revision, they may have put less mental effort in processing teacher feedback.

#### 2.4.2.3 *Summary of findings on processing of feedback in revision*

Available peer feedback and teacher feedback studies have identified cognitive and metacognitive operations that L2 learners conducted when processing feedback during revision of their drafts. These included employing symbols and external resources, using L1, activating previous knowledge, reflecting, providing metalinguistic explanations, predicting, making connections, planning, managing attention, monitoring and evaluating.

None of existing peer feedback studies, however, have incorporated think-aloud protocols (see L. J. Zhang & Cheng, 2020). Since the think-aloud method may “help researchers move closer to opening the ‘black box’ that lies between peer feedback as input and improvement on writing as output” (Hu & Lam, 2010, p. 390), there exists the need to incorporate think-aloud methods in peer feedback studies.



## **2.5 Student perceptions of peer feedback**

Both L1 and L2 studies have investigated into students' perceptions of peer feedback. These studies have elicited data from interviews, reflective essays, and questionnaires. As the findings from L1 and L2 peer feedback studies are disparate, they are reviewed separately.

### **2.5.1 L1 learners' perceptions of peer feedback**

Because there are few L1 peer feedback studies, the findings on L1 learners' perceptions of peer feedback are also limited. Quite unexpectedly, L1 learners demonstrated contrasting attitudes towards peer feedback.

Only two studies (Cho, Schunn, & Charney, 2006; Katstra, Tollefson, & Gilbert, 1987) reported that their participants were positive about peer feedback. Using questionnaires with 177 participants before and after the intervention, Katstra et al. (1987) found that their participants assumed that peer feedback helped them make improvements in writing. In addition, they expressed enthusiasm for some peer feedback activities, including sharing their writing with their peers, reading their peers' compositions and giving feedback, and rewriting based on feedback that they received. In Cho, Schunn, and Charney's (2006) study, the researchers purposely kept their participants from knowing the feedback source and found that their participants perceived feedback from peers and from experts to be equally helpful.

However, when learners are aware that they receive feedback from their peers, their perceptions are not so positive. The participants in Gielen et al.'s (2010) study reacted extremely negatively to peer feedback and very few of them expressed a desire to use it in future assignments or courses. Cho, Schunn & Wilson (2006) reported that the

participants regarded peer feedback as having low validity and reliability. Similarly, Kaufman and Schunn (2011) found that their participants, who only received peer feedback, regarded peer feedback as unfair, a perception that dropped significantly over time. Although it has been posited that learners' negative perceptions do not threaten the effectiveness of peer feedback (Gielen et al., 2010; Kaufman & Schunn, 2011), such perceptions need to be addressed. When not addressed, they may affect learners' participation in peer feedback activities and therefore compromise the effects of peer feedback.

### **2.5.2 L2 learners' perceptions of peer feedback**

Unlike the limited findings from L1 peer feedback studies, there are substantial findings about L2 learners' perceptions of peer feedback. Findings from these studies are mixed yet encouraging.

Some studies reported L2 learners' positive attitudes towards peer feedback. van den Berg et al. (2006), based on questionnaire data, found that 80% of the 131 participants who participated in peer feedback activities over one semester attributed their progress in writing to peer feedback. Similarly, J. Zhou et al. (2020) reported that of the 30 posts recording the students' feelings about peer feedback, only two conveyed a negative attitude. In other studies, L2 learners have reported that they learned from others' strong points to offset their own weaknesses, benefited from collaborative learning opportunities afforded by peer feedback, obtained multiple viewpoints offered by their peers, enhanced audience awareness, as well as developed critical thinking abilities (Hu & Lam, 2010; Mendonça & Johnson, 1994; Ruegg, 2017; Tsui & Ng, 2000; M. Yang et al., 2006). Min (2008), by interviewing the 18 participants, found a qualitative change in their perceptions of peer feedback. Initially, most of the participants reported that peer

feedback was delivered in a commanding tone, vague and unhelpful; after intensive and systematic peer feedback training and four writing tasks with peer feedback activities, they expressed their satisfaction with thought-provoking and specific feedback from their peers that were collaborative in tone.

Several studies have reported L2 learners' preference for teacher feedback over peer feedback, as well as their concerns about peer feedback. Students' preference for teacher feedback has been attributed to multiple factors: the cultural beliefs and attitudes (viewing teachers as authorities), students' trust in the expertise possessed by teachers, and their familiarity with the teacher feedback practice (Hu & Lam, 2010; Lee, 2015; Nelson & Carson, 2006; Zhao, 2010). Meanwhile, a variety of factors affecting L2 learners' perceived helpfulness of peer feedback have been identified. These included: the limited language proficiency and knowledge of L2 learners (Liou & Peng, 2009; Mangelsdorf & Schlumberger, 1992; Poverjuc, Brooks, & Wray, 2012; Tsui & Ng, 2000; Y. Yang & Meng, 2013), their peers' attitudes and stances in peer feedback activities (McGroarty & Zhu, 1997; J. Zhou et al., 2020), peer feedback quality (Lee, 2015; Mangelsdorf, 1992; Min, 2005; Nelson & Carson, 1998; Ruegg, 2017), and time constraints for in-class peer feedback activities (W. Wang, 2014; Q. Zhu, 2018).

Sengupta (1998) conducted a study in a Hong Kong secondary school, reporting that the EFL learners' negative attitude towards peer feedback. In the interviews, the six participants said nothing positive about peer feedback: They voiced several concerns about peer feedback, perceived their teachers to be the only reader, and considered them and their peers inadequate for completing peer feedback tasks. In her study, Sengupta (1998) attributed the participants' negative perceptions of peer feedback to the "examination-driven" and "accuracy-oriented" educational context in Hong Kong and

the lack of instruction in peer evaluation for students (p. 19). These two factors may also help explain the different finding in Sengupta (1998) from other L2 peer feedback studies.

As well as reporting general perceptions of peer feedback, some studies have touched on L2 learners' perceptions of specific elements of the peer feedback intervention. Tsui and Ng (2000) found Chinese EFL learners appreciated reading their peers' compositions. Using a six-point Likert scale to measure the perceived helpfulness of giving and receiving feedback, the researchers found the overall means for the usefulness of reading peers' compositions was 4.14, while the overall means for the usefulness of peers' oral comments and for the usefulness of peers' written comments received were 3.54 and 3.64, respectively. The significant difference in the scores indicated a higher degree of perceived helpfulness of giving feedback than of receiving feedback. Similarly, other studies (e.g., van den Berg et al., 2006; Ruegg, 2017; J. Zhou et al., 2020) have reported that L2 students were enthusiastic about reading and evaluating their peers' compositions. They acknowledged that from reading their peers' compositions, they learned new ideas and writing strategies and improved their own writing.

Of the participants in van den Berg et al.'s (2006) study, 83% acknowledged that they valued the received written feedback and 78% reported that they valued the received oral feedback. In addition, L2 students expressed a preference for negative feedback which led to changes and consequently improved their drafts, but they reported that feedback on grammar and sentence level details was ineffective (Nelson & Carson, 1998).

Some studies reported L2 learners' perceptions of using scoring rubrics and incorporating oral interactions in peer feedback activities. While L2 students felt that the rubrics made the criteria of successful EFL writing transparent and facilitated the provision of feedback (W. Wang, 2014; J. Zhou et al., 2020), they expressed their concerns about using the same rubric throughout the intervention, as it could result in instrumentalism, "leading students to write to the rubric to achieve a high score" (W. Wang, 2014, p. 93). For oral interactions, L2 learners expressed their appreciation of opportunities of collaborative learning, "to clarify their intended meanings to the reader and to negotiate a way to convey the intended meaning more effectively" (Tsui & Ng, 2000, p. 166), but also they expressed their need for more time for oral interactions (Q. Zhu, 2018).

### **2.5.3 Summary of findings on students' perceptions of peer feedback**

L1 and L2 studies reported that students displayed mixed attitudes towards peer feedback. Some studies documented students' positive attitude (e.g., Cho, Schunn, & Charney, 2006; Katstra, Tollefson, & Gilbert, 1987; van den Berg et al., 2006; J. Zhou et al., 2020); some studies reported student' preference of teacher feedback over peer feedback and their concerns for peer feedback (e.g., Hu & Lam, 2010; Tsui & Ng, 2000; M. Yang et al., 2006); and other studies showed students' negative attitudes (e.g., Cho, Schunn & Wilson, 2006; Gielen et al., 2010; Kaufman & Schunn, 2011; Sengupta, 1998).

Compared with studies reporting learners' general attitudes of peer feedback, the studies examining learners' perceptions of specific elements within the peer feedback intervention have been limited in number. No studies to date have examined learners' perceptions of peer feedback training, although there have been numerous empirical and

descriptive studies on peer feedback training (e.g., Hu, 2005; Lam, 2010; McGroarty & Zhu, 1997; Min, 2005, 2006, 2008; Rahimi, 2013). Similarly, few studies have tapped into learners' perceptions of oral interactions which were extensively researched in terms of use of communication strategies and interaction patterns (Mangelsdorf & Schlumberger, 1992; Mendonça & Johnson, 1994; Nelson & Murphy, 1993; Zhao, 2018; W. Zhu & Mitchell, 2012). The exploration into learners' perceptions of specific elements in the peer feedback intervention could corroborate the quantitative findings for an intervention and provide a rich source of evidence for how, and why, an intervention is effective or not, thus offering valuable insights on how the intervention could be optimized.

The above-mentioned studies also have a methodological limitation that might have threatened validity of the findings. In some previous studies (Hu & Lam, 2010; Tsui & Ng, 2000; Zhao, 2010), the perception data of peer feedback were collected when students had experienced both teacher feedback and peer feedback. The procedure of implementing teacher feedback alongside peer feedback was "pedagogically rational" in writing instruction, because teachers have the responsibility to monitor students' learning (Lee, 2015, p. 3). Yet, in a research setting that looks specifically into students' perceptions of peer feedback, the influence of teacher feedback on peer feedback should be minimized. Researchers may therefore consider keeping the participants in the treatment group from receiving other forms of feedback, as Ruegg (2017) did.

## **2.6 Conceptual framework for the present study**

The social constructivist theory, traditionally associated with Vygotsky's (1978) work, contends that cognitive development and knowledge is socially constructed through collaboration, interaction and communication with others. Such a perspective views

learning as occurring in social interactions, rather than within an individual. For the present study, its three concepts, mediation, zone of proximal development (ZPD) and scaffolding guide the design of the peer feedback intervention.

Mediation, the relationship that humans have established with the world by relying on physical tools and psychological tools, is the most fundamental concept in social constructivism (Lantolf, 2000). Human mind is mediated and learning is a mental process that requires mediation (Vygotsky, 1978). The concept of mediation is elaborated and developed in other works. Lantolf (2000) distinguished three forms of mediation: social mediation (interactions with others), self-mediation (self-directed speech), and artefact mediation (such as portfolios, tasks, and technology), the majority of which are heavily dependent on “psychological tools such as language, signs, and symbols” (Karpov & Haywood, 1998, p. 27). The activity theory, proposed by Leontiev and further developed by Engeström, is also based upon the concept of mediation.

The ZPD and scaffolding are two related concepts in social constructivism, although only the former one was originally proposed by Vygotsky. ZPD is defined as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). It is viewed as an optimal scenario for learning in which a learner receives assistance from more competent others while engaging in learning activities. Wood, Bruner and Ross (1976) defined such assistance as scaffolding and identified six features characterizing scaffolding,

[S]caffolding help was characterized by several features which include encouraging the learner to take an interest in the task [recruitment], simplifying

the task [reduction in degrees of freedom] and maintaining interest in its objectives [direction maintenance], controlling and reducing frustration during the activity [frustration control], focusing the learner on critical features of the task [marking critical features], and providing demonstrations of how to achieve the goals of the task [demonstration]. (Wigglesworth & Storch, 2012, p. 72)

Other researchers, based on their empirical findings that collaborative work among learners also provided opportunities for scaffolded help, extended the concept of scaffolding by including novice-novice interactions (de Guerrero & Villamil, 1994, 2000; Swain & Lapkin, 1998; Villamil & de Guerrero, 1996, 1998). Donato (1994) put forward the concept of collective scaffolding, occurring routinely between or among peers engaging in language learning. He justified the concept based on ZPD arguing that all learners had strengths and weaknesses that emerge in peer interactions, and that they were able to pool their resources and scaffold each other's learning (Donato, 1994).

From the perspective of social constructivism, any teaching activity that hopes to develop learning should situate learners in social interactions, in which they can receive assistance from an expert or work collaboratively with their peers. Many researchers have adopted sociocultural theory to analyze positive effects that social interaction has on second or foreign language learning (Aljaafreh & Lantolf, 1994; de Guerrero & Villamil, 2000; Donato, 1994; H. H. Li, Zhang, & Parr, 2020; Ohta, 2000; Storch, 1999; Swain, 2000; Wigglesworth & Storch, 2012). For instance, Aljaafreh and Lantolf (1994) claimed instructional conversation and scaffolding from the teacher within the students' ZPD led to L2 language learning. Other studies have examined interactions between peers. Ohta (2000) examined interactions between two Japanese learners working on a translation task and found the weaker learner was able to accomplish the translation task



fluently and accurately with support from the peer. Swain (2000) argued that language related episodes, which occurred in peer interactions, were occasions for language learning and such learning can be transformed into “individual mental resources” (p.104). De Guerrero and Villamil (2000) adopted a microgenetic approach to analyze the interactions between a reviewer and a writer in peer revision. The writer was found to incorporate the majority of changes discussed and, in some instances, to further revise on his own, while the reviewer also made progress in writing, in revising, and in his/her ability to provide assistance. The pair work “allowed both reader and writer to consolidate and reorganize knowledge of the L2 in structural and rhetorical aspects and to make his knowledge for each other’s benefits” (de Guerrero & Villamil, 2000, p. 65). Similarly, Wigglesworth and Storch (2012) argued that pair work was beneficial for writing development in that it created learning opportunities which “can allow students to scaffold each other’s contributions, and to pool their resources and discuss, and thus engage with, the feedback they are writing” (p.92). A recent study, H. H. Li et al. (2020), conducted in the Chinese EFL context, found peer talk before individual writing fulfilled four functions: enabling learners to generate content, organization and language for their individual writing, providing learners with opportunities of collaboratively solving linguistic problems, organizing and managing the group task, and sharing emotions and maintaining group harmony.

In the present study, the dyadic peer feedback activities situated learners in peer-mediated social interactions in which they collaborated with their peers, exchanging written and oral feedback. The exchange of written feedback, and oral feedback in particular, provided learners with learning opportunities because they could pool their strengths and weaknesses, negotiate feedback, and solve problems in their drafts together.

Also, the concepts of ZPD and scaffolding guide the design of peer feedback intervention in the present study. Because all participants in this study had similar language proficiency, being categorized as immediate-advanced learners based on their scores in the placement test, they could be assumed to have similar ZPDs, and thus be capable of learning from each other. Peer scaffolding occurred when students exchanged their written feedback points and when they orally discussed the feedback with their peers. In addition, teacher scaffolding was implemented in this study in a sustained manner, including a general training session, six focused training sessions and corresponding peer feedback forms.

## **2.7 Chapter summary**

This chapter started with an introduction to the process approach to writing and the theories that inform peer feedback. It then critically reviewed both L1 and L2 feedback studies relevant to the foci of the present study, the impact of peer feedback, processing of peer feedback in revision, and student perceptions of peer feedback. Finally, it presented the conceptual framework for the design of the intervention in the present study. The next chapter will discuss the design and methodology of this study.

## **Chapter 3 Methodology**

### **3.1 Chapter overview**

The chapter describes first the research design and instructional context for this study and is followed by the preparatory stage and the procedures. The three parts of the study are then described in detail, including the participants, data collection procedures, and data analysis procedures. The chapter ends with a brief discussion of ethical considerations.

### **3.2 Overview of research design**

#### **3.2.1 Research paradigms and approaches**

Paradigms, also known as worldviews and knowledge claims, are philosophical ideas that guide a researcher's inquiry. There are four paradigms: the post-positivism, constructivism, the transformative world view, and pragmatism.

Post-positivism underlies quantitative methods, relying on "careful observation and measurement of the objective reality" to test, or verify and refine theories (Creswell & Creswell, 2018, p. 7). By collecting numerical data from participants, researchers generalize, validate hypotheses, or construct theories.

Both constructivism and the transformative worldview advocate for qualitative research. Constructivism holds the following principle: Human beings construct meanings, which are subjective, socially and historically negotiated, as they engage with the world. Constructivist researchers acknowledge that their interpretation of meaning is shaped by their own experience and backgrounds, and they use open-ended questions to collect data from the field, interpret what they find, and construct meanings inductively. A

transformative worldview advocates for “an action agenda” to emancipate marginalized people from constraints (Creswell & Creswell, 2018, p. 9). Researchers work collaboratively with the participants to bring changes to the participants.

Pragmatism is a practice-driven approach which emphasizes research questions and advocates for the employment of pluralistic methods such as experiments, case studies, and surveys. Mixed-methods research, therefore, is premised on pragmatism.

### 3.2.2 Research design

This study, guided by the pragmatic worldview, employed a convergent design (Creswell & Creswell, 2018); data collected from different sources were integrated when the research results were interpreted (see Figure 3.1).

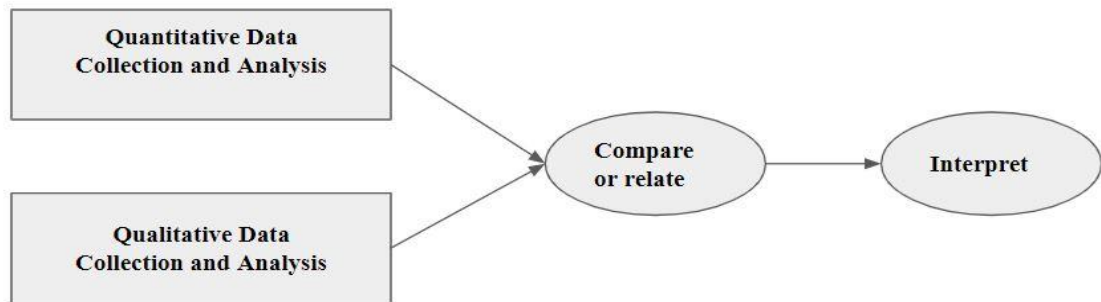


Figure 3.1 Convergent design

Data for this study were from three sources: Product data (quantitative data), process data (qualitative data) and perception data (qualitative data). Part 1, a quasi-

experimental study, examined the effects of peer feedback on text revisions and on writing performance by comparing the participants' two drafts of one writing task and their compositions in the pre-test, post-test, and delayed post-test (quantitative data). Part 2 was a case study which explored how the treatment group participants made use of peer feedback when writing their second drafts (qualitative data); the findings cross-checked the quantitative findings about the effects of peer feedback on text revisions. In Part 3, questionnaire and interviews were used to investigate the treatment group participants' perceptions of peer feedback (qualitative data); the findings cross-checked the quantitative findings about the effects of peer feedback on writing performance. An overview of the research design is shown in Table 3.1.

Table 3.1 An overview of the research design

	<b>Research Objectives</b>	<b>Data</b>	<b>Participants</b>
Part 1	The effects of peer feedback on text revisions	<ul style="list-style-type: none"> <li>• Two drafts (Draft 1 and Draft 2) of Writing Task 6</li> </ul>	Treatment group Comparison group
	The effects of peer feedback on writing performance	<ul style="list-style-type: none"> <li>• Compositions from the pre-, post-, and delayed post-tests</li> </ul>	
Part 2	Processing of peer feedback in revision	<ul style="list-style-type: none"> <li>• Written documents in Writing Task 6, including:                             <ul style="list-style-type: none"> <li>-Draft 1 with written feedback,</li> <li>-Peer feedback form (PF form),</li> <li>-Draft 2</li> </ul> </li> </ul>	Six participants from the treatment group
		<ul style="list-style-type: none"> <li>• Transcripts of paired discussions in Writing Task 6</li> </ul>	
		<ul style="list-style-type: none"> <li>• Transcripts of think-aloud protocols in Writing Task 6</li> </ul>	
Part 3	Perceptions of peer feedback	<ul style="list-style-type: none"> <li>• Percentages and quotes from the Perceptions of Peer Feedback Questionnaire</li> </ul>	Treatment group
		<ul style="list-style-type: none"> <li>• Transcripts of semi-structured interviews</li> </ul>	The six case study participants

### **3.3 Instructional context**

The University (University X), where the present study was conducted, is a national key university, under the direct jurisdiction of the Ministry of Education. University X is a member university of “Project 211” and “Project 985”, two prestigious national projects supporting the development of high-quality, world-class universities in China.

In University X, English courses for non-English majors at the undergraduate level include College English as a compulsory course, English electives, and major-related English for Specific Purposes (ESP) courses. All non-English majors, on entering the university, participate in a placement test, which classifies them into Level One and Level Two in the order of ascending proficiency, based on their scores. Typically, 80% start with Level One and 20% start with Level Two. All non-English majors, irrespective of the level at which they start, are required to take the College English course in the first two years (four semesters in total). For the College English course in University X, a variety of textbooks are used to meet the needs of learners at different proficiency levels and the language requirements of different majors and disciplines. University X has also purchased a variety of online resources and encouraged teachers and students to use them in language teaching and learning.

The present study took place in a 17-week College English course (Level Three) in the second semester of the first academic year. In the 17-week course, the students in two intact classes were required to write six compositions (each in two drafts) and participate in feedback activities. Their performance in the six writing tasks and feedback activities contributed to 10% of their total score for the course.

### **3.4 Preparatory stage: Instrument and procedure trial and modification**

It is regarded as a good practice for researchers to pilot their instruments and research procedures before launching their projects. The preparatory stage, in which the validity of instruments and procedures of data collection and analysis are assessed, is essential for both quantitative and qualitative studies (Creswell & Creswell, 2018). Based on the findings from this stage, researchers usually make modifications to their instruments and procedures in order to avoid “a great deal of frustration and possible extra work” in later stages of their research (Dörnyei, 2007, p. 75). I followed these suggestions in the process of executing my research.

One month prior to this study, two instruments, the Perceptions of Peer Feedback Questionnaire and Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey’s (1981) ESL Composition Profile, and a writing task were piloted. The participants were four second-year non-English major students who had peer feedback experience through taking the elective course of Academic Writing, and two EFL teachers. Both EFL teachers had a master’s degree in Applied Linguistics and more than 10 years of teaching experience.

#### **3.4.1 Instrument refinement**

The four students were invited to complete the Perceptions of Peer Feedback Questionnaire within 30 minutes at a meeting on campus. Because the main objective of this stage was to refine instruments and improve procedures, the questionnaires were not collected for statistical analysis. After completing the questionnaire, the four students were invited to reread the questionnaire to comment on clarity and comprehensibility of instructions and time allowance for the questionnaire.

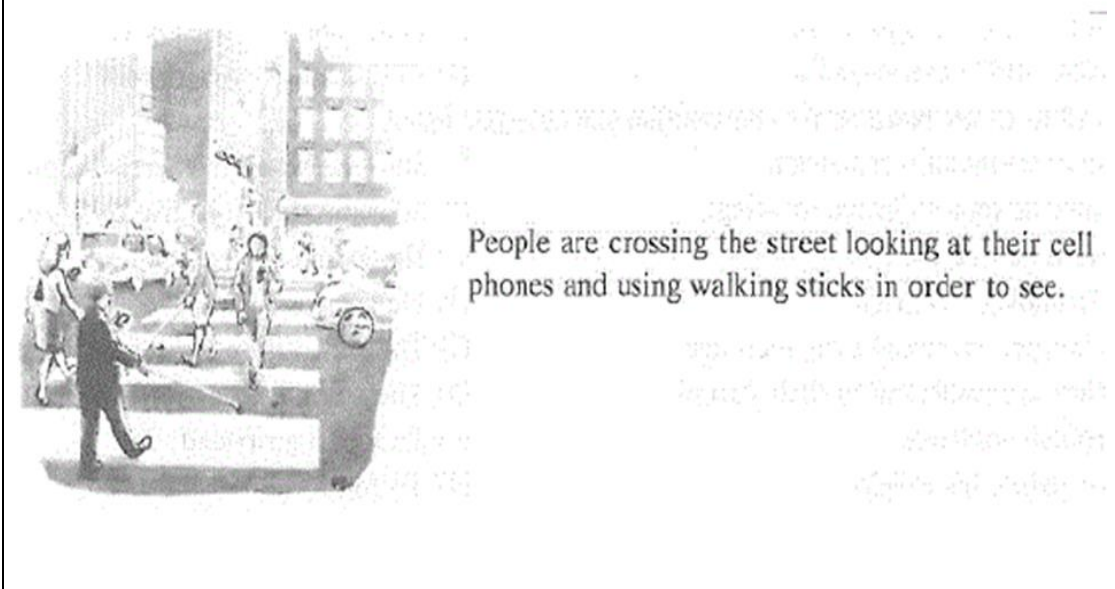


The two EFL teachers were invited to read through and comment on Jacobs et al.'s (1981) ESL Composition Profile, which has been used successfully to evaluate writing proficiency of ESL and EFL students in various contexts. They identified two problems in the scoring rubric that could impede the reliability of scoring. Firstly, within the five constructs (content, organization, vocabulary, language use, and mechanics) of the scoring rubric, "succinct" and "limited support" in the organization construct overlapped with "thorough development of thesis" and "limited development of thesis" in the content construct. Jacobs et al. (1981) explained that the criterion for evaluating "thorough development of thesis" is "Is the thesis expanded enough to convey a sense of completeness?" (p.92), and the criterion for "succinct" is "Are all ideas directed concisely to the central focus of the paper, without digression?" (p.93). Secondly, the two EFL teachers disagreed with paragraphing being a descriptor for mechanics, "Are paragraphs indented to indicate when one sequence of thought ends and another begins?" (Jacobs et al., 1981, p. 96). I, therefore, revised the scoring rubric accordingly based on the two teachers' opinions.

#### **3.4.2 Intervention procedure trial**

The four students participated in a writing task with peer feedback activities. Firstly, they were asked to complete a writing task on site, with a topic selected from previous CET-4 writing tests, for which the time allowance and word number had been modified. The prompts are as follows.

**Directions:** Write a short essay based on the picture below. You should start your essay with a brief account of the increasing use of the mobile phone in people's life and then explain the consequences of overusing it. You should write at least 200 words.



The four students then formed two pairs. Working in pairs, they exchanged their compositions. They were asked to provide written feedback for their peers, using a PF form and relying on whatever resources they have access to. Although I did not set a time limit for this activity, all four students finished the task of providing written feedback within forty minutes. A 20-minute paired discussion session followed which I observed but did not participate in.

Finally, the four students worked independently on their second drafts based on peer feedback and their discussions. Two students returned their second drafts after approximately 45 minutes and the other two after 60 minutes. The two drafts written by the four students were not collected.

Immediately after the intervention trial, I had a group discussion with the four students, inviting them to comment on the tasks in terms of comprehensibility and manageability. The four students indicated that they could accomplish the tasks as required, but would like to have more time for the paired-discussion.

### **3.4.3 Some changes to the instruments and procedures**

Based on comments from the two teachers and the four students, and my observation, some changes were made to the implementation of the research design.

Firstly, time limits for some activities were adjusted. I decided to extend the time limit for the pre-test, post-test, and delayed post-test to forty-five minutes, considering the proficiency of the participants in this study. I also extended the time limit for discussion to thirty minutes, based on feedback from the four students and my discussion with the teacher participant in this study.

Secondly, some modifications were made to the intervention procedure to improve the quality of data collection. Given the duration of the study and time constraints for the College English course, the teacher participant and I jointly decided that only some activities (peer feedback training and paired discussions) should be done during class time and other activities (writing the first draft, providing written comments for peers, and writing the second draft) should be done as after-class assignments. In this way, the time could be controlled, and potentially confounding variables would remain minimal.

Thirdly, Jacobs et al.'s (1981) ESL Composition Profile was amended to improve construct validity and clarity. To be specific, "succinct" and "limited support" were deleted from the organization construct; "paragraphing" in the mechanics construct was replaced by "indentation/extra-spacing for paragraphing". In addition, the construct

“language use” was renamed “grammar”, since the former was more appropriate as a general term including grammar, vocabulary, and mechanics. The criteria for handwriting in the construct of mechanics were deleted as the raters marked the compositions in electronic format. The adapted Jacobs et al.’s (1981) ESL Composition Profile is in Appendix A.

### 3.5 Procedures for the study

The study included a pre-test, an intervention for 13 weeks, a post-test, a post-intervention for 6 weeks, and a delayed post-test (see Table 3.2). During the 13-week intervention, the treatment group participated in peer feedback training and other peer feedback activities, while the comparison group received collective feedback regularly.

Table 3.2 Procedures for the study

<b>Time</b>	<b>Procedures</b>
Week 1	Pre-test
Week 2	Peer feedback training for the treatment group
Week 3-14	Six writing tasks with peer/collective feedback activities <u>Case study participants:</u> Think-aloud training (Week 11) and think-aloud session (Week 13)
Week 16	Post-test  <u>Treatment group:</u> The Perceptions of Peer Feedback Questionnaire <u>Case study participants:</u> Semi-structured interviews
Week 19-24	Three independent writing tasks
Week 25	Delayed post-test

### 3.5.1 Pre-test, post-test, and delayed post-test weeks

The pre-test, post-test and delayed post-test were administered in class in Week 1, Week 16, and Week 25 respectively. Participants in the treatment group and comparison group took the three tests under the same conditions. As will be discussed in 3.6.2, the writing topics for the three tests were different, but they were of the same difficulty level. In the post-test week, the Perceptions of Peer Feedback Questionnaire was administered to the treatment group only.

### 3.5.2 Peer feedback training

As discussed in 2.3.3, peer feedback training is crucial for the success of peer feedback intervention. Training, in this study, was offered through a general training session and six focused training sessions. I took the general training, while the focused training was taken by the teacher. The general training, two sessions of the College English course which took 90 minutes, was conducted in the second week, prior to the six writing tasks. The procedures and activities for the general training were adapted from Bui and Kong (2019), Hu (2005), Lam (2010) and Min (2005), as summarized in Table 3.3.

Table 3.3 Stages and activities for the general training session

<b>Stages</b>	<b>Activities</b>
Awareness-raising	<ul style="list-style-type: none"><li>• Two video clips of peer feedback activities</li><li>• Group discussions about peer feedback</li><li>• Presentation: An introduction to peer feedback</li></ul>
Instruction & Demonstration	<ul style="list-style-type: none"><li>• Instruction on the four-step procedure in generating feedback</li><li>• Demonstration of reviewing a sample paragraph</li><li>• Demonstration of revising a sample paragraph</li></ul>

Practice	<ul style="list-style-type: none"> <li>• Whole-class providing feedback to a sample paragraph</li> <li>• Whole-class discussion about the appropriateness of the feedback</li> <li>• Each student independently revising the sample paragraph based on the feedback</li> </ul>
Reflection & review	<ul style="list-style-type: none"> <li>• Students raising questions about peer feedback</li> <li>• Summary of key concepts of peer feedback</li> </ul>

There were four stages in the general training: Awareness-raising, demonstration, practice, and reflection & review. In the awareness-raising stage, students watched two video clips of peer feedback activities and then discussed in groups, to develop a basic understanding about what peer feedback was and how it would be implemented. A systematic introduction to peer feedback was then presented, focusing on the following topics: the nature of peer feedback, its contribution to writing development, priorities in giving and receiving feedback, and stances in peer feedback activities.

In the instruction & demonstration stage, the four-step procedure in providing feedback, which includes “clarifying writers’ intentions”, “identifying problems”, “explaining the nature of problems”, and “making specific suggestions” (Min, 2005, p.293), was first explained. What followed was a demonstration of the reviewing process using the four-step procedure and a PF form, and the revising process with feedback.

The practice stage started with the students reviewing another sample paragraph. They read and commented on the sample paragraph as a whole class, followed by a teacher-led discussion about the appropriateness of their feedback. They then individually revised the sample paragraph by incorporating the feedback.

In the final stage, through group and class discussions, the students shared difficulties they encountered when giving feedback and making revisions. The general training session ended with a brief review of key concepts related to peer feedback.

In addition, the six focused training sessions, taking place within the six writing tasks, featured specific topics (see Table 3.4). Within each session, the teacher, by using examples from students' drafts and PF forms, compared between useful and problematic feedback points and gave suggestions on how to use them in revision. Occasionally, the teacher addressed other common problems she identified from the treatment group participants' drafts.

Table 3.4 Topics in the six focused peer training sessions

<b>Writing task</b>	<b>Focus</b>
<b>1</b>	Useful and problematic feedback & revisions in response to feedback
<b>2</b>	Evaluating and revising thesis statement
<b>3</b>	Evaluating and revising body paragraphs
<b>4</b>	Evaluating and revising language problems: Grammar and sentence structures
<b>5</b>	Evaluating and revising language problems: Vocabulary Using reference resources: Dictionaries, corpus, scoring rubric and exemplar essays
<b>6</b>	Evaluating and revising problems in cohesion

Take the focused training session for Writing Task 2 as an example. The session started with a review of the three tips for writing a thesis statement: “A thesis statement is an assertion, not a fact”; “A thesis statement cannot be framed into a question”; “A thesis

statement should be concise (using one sentence) and summative (covering arguments in the three body paragraphs)” (PowerPoint slides for Writing Task 2). The teacher then presented some helpful comments and some problematic ones on thesis statement selected from the students’ drafts and PF forms, and asked the students to analyze them in the context. Also, the teacher reminded the students to make judicious use of the comments they received and make self-revisions when needed. As the students had not reached the stage of writing their second drafts, the teacher gave suggestions but refrained herself from providing revision solutions.

### **3.5.3 Six writing tasks with peer feedback activities**

The six writing tasks started in Week Three, with each writing task taking four sessions over two weeks. With each writing task, the treatment group participants wrote their first drafts, exchanged their drafts in pairs, provided each other with written feedback, participated in a focused peer feedback training session and a paired discussion session about written feedback, and finally wrote their second drafts. All the six writing tasks involved expository topics, selected from previous CET-4 writing tests (see Appendix B).

With peer feedback activities, the treatment group participants worked in pairs but their partners changed for each writing task. In addition, PF forms with varying foci (in accordance with the topics in focused training sessions) were used for the written feedback (see Appendix C). When providing written feedback for their peers, the participants were required to follow and complete a PF form; they were told that in-text or end-of-text comments could also be provided on their peers’ first drafts.

While the treatment group participants engaged in writing tasks with peer feedback activities, the comparison group participants wrote on the same topic and received



collective feedback. As explained in 1.4, the teacher typically provided feedback for 1/2 to 1/3 of the compositions collected from the comparison group participants, summarized common problems and finally presented them to the whole class. The comparison group participants wrote their second drafts based on the collective feedback.

At the close of each writing task, both groups were required to hand in their written documents. Specifically, the treatment group participants handed in their first drafts, PF forms, and second drafts; while the comparison group participants handed in their first and second drafts.

#### **3.5.4 Three independent writing tasks**

The three independent writing tasks, for which neither group received feedback, started in Week 19. The topics were expository, selected from previous CET-4 writing tests, and did not overlap with the six topics in the intervention (see Appendix D).

As the three independent writing tasks occurred during the summer vacation, participants in both groups were asked to hand in their compositions through Pigaiwang, an automated online writing evaluation system. On Monday of Week 19, Week 21, and Week 23, the teacher assigned topics through Pigaiwang; on Friday of Week 20, Week 22, and Week 24, participants in both groups were asked to upload their final drafts. The teacher suggested that students write their first drafts within forty-five minutes, and make revisions, using whatever resources available before uploading their final drafts. The teacher made it clear to both groups that no feedback would be offered for the three independent writing tasks.

### **3.6 Part One: Effects of peer feedback on text revisions and writing performance**

A quasi-experimental design was used in the first part of this study with two conditions, a treatment group and a comparison group, to address the following two research questions and their sub-questions.

RQ 1 How does peer feedback influence Chinese EFL learners' text revisions?

*How does peer feedback influence the overall quality, content quality and organization quality of Chinese EFL learners' text revisions?*

*How does peer feedback influence the accuracy, syntactic complexity, lexical complexity, and fluency of Chinese EFL learners' text revisions?*

RQ 2 How does peer feedback influence Chinese EFL learners' writing performance?

*How does peer feedback influence Chinese EFL learners' writing performance in terms of the overall quality, content quality and organization quality?*

*How does peer feedback influence Chinese EFL learners' writing performance in terms of complexity, accuracy, and fluency?*

A quasi-experimental design, as one type of experimental methods (with the other three being pre-experimental design, single-subject design, and true experiment), suits the present study best. Pre-experimental and single-subject experimental methods are not appropriate in that pre-experimental designs include a single group, and therefore having no control group to compare with the experimental group, and single-subject designs involve over-time observation of an individual (Creswell & Creswell, 2018). The experimental design, the most advantageous of the four types because of random assignment of participants to treatment conditions, however, is not practical for research

with participants in naturally-formed groups, such as classrooms and organizations (Muijs, 2010). The quasi-experimental design, including a treatment group and a comparison group without random assignment, is a practical research approach (Dörnyei, 2007). For this reason, I decided to adopt a quasi-experimental design.

### 3.6.1 Participants

Participants in the quasi-experimental study were the students in the two classes taught by a female Chinese EFL teacher in University X. Convenience sampling was used for the quasi-experimental study, because it is a practical way to recruit participants (Creswell & Creswell, 2018). Even though this sampling method has been criticized for not representing the population (Dörnyei, 2007), this limitation may not be a problem in this study, as students within each class were from different majors. Also, when organizing College English courses, the university gave consideration to the number of students, gender distribution, and average grade for each class to ensure all classes of the same level are analogous.

The teacher participant decided that the class with 60 students be the treatment group and the class with 45 students be the comparison group. Of the 105 students in the two classes, 47 students in the treatment group and 37 students in the comparison group participated in the three tests (pre-test, post-test, and delayed post-test) and completed the six writing tasks with feedback activities and the three independent writing tasks. The data of six students in the treatment group, who participated in the case study, were removed from the quasi-experimental study. Table 3.5 presents the number of participants in the quasi-experimental study.

Table 3.5 Participants in the quasi-experimental study

<b>Treatment group</b>	N=41
<b>Comparison group</b>	N=37

All participants are of Chinese nationality and from science majors, Electrical Science, Biology, Chemistry, or Physics, and were in the second semester of their first academic year. They were classified as upper-intermediate EFL learners, whose language proficiency was equivalent to IELTS 5.5 or 6.0 (Common European Framework of Reference B2 equivalency), as assessed by the university English placement test.

To facilitate the analysis of data and reporting of findings, each participant in the quasi-experimental study was assigned a number as a unique identification code and was required to use it consistently in the three tests, in the six writing tasks, as well in their questionnaires. Specifically, 1-41 were assigned to the treatment group participants and 42-78 were assigned to the comparison group participants.

### 3.6.2 Data collection

Data from the quasi-experimental study included compositions written by the treatment group and comparison group participants. To examine the effects of peer feedback on text revisions, two drafts of Writing Task 6 were collected from both groups; to investigate the effects of peer feedback on writing performance, compositions in the pre-test, post-test, and delayed post-test were collected. In total, 390 compositions were collected (see Table 3.6).

Table 3.6 Compositions collected for the quasi-experimental study

	<b>Treatment group</b>	<b>Comparison group</b>
Draft 1 (Writing Task 6)	41	37
Draft 2 (Writing Task 6)	41	37
Pre-test	41	37
Post-test	41	37
Delayed post-test	41	37
Total		390

Topics for Writing Task 6, and for the three writing tests, were expository topics chosen from previous CET-4 writing tests because of the status of CET within the curriculum as well as the role of expository essays in the writing instruction (see Appendix B and Appendix E).

CET, designed and implemented by the Ministry of Education of the People’s Republic of China, is a standardized test designed to evaluate the English proficiency of non-English majors in China. CET has been the most influential test in China for more than 30 years and, in 2006, it tested 12,000,000 Chinese EFL learners (H. Li, Zhong, & Suen, 2012). After many rounds of reforms, the current CET exhibits “acceptable levels of reliability and validity” (Ying Zheng & Cheng, 2008, p.414).

The choice of expository topics was decided through taking account of the frequency of occurrence of different genres in L1 and L2 academic settings. An expository essay is defined as “a genre through which writers present a point of view and support it with examples and evidence” (Schleppegrell, 2004, p. 88). In L1 contexts, expository essays have been identified as “the most prevalent” and “the most valued” genre in secondary

school instruction and assessment (Llosa, Beck, & Zhao, 2011, p. 256), as well as one of the “the most frequent essay genres” that university students write (Gardner & Nesi, 2013, p. 42). In EFL contexts, expository topics are also frequently used in domestic English tests for English majors and non-English majors, and in international English tests such as the International English Language Testing Systems (IELTS) and the Test of English as a Foreign Language (TOEFL).

The teacher and I jointly made two decisions about data collection. Firstly, three parallel topics from one CET-4 writing test were chosen for the three writing tests to ensure that the topics were at the same difficulty level. Parallel topics within one CET test were more likely to be equal in difficulty level in comparison with personal judgment of difficulty level. The second decision was that the three writing tests would be conducted in class to control other variables that might influence participants’ writing performance, such as seeking help from others and referring to the internet or textbooks for information. The two drafts of Writing Task 6 were completed as after-class assignments, and were collected at the end of Writing Task 6.

### **3.6.3 Data analysis**

Analysis of the data from the quasi-experimental study was conducted in two steps: scoring and coding the participants’ compositions and statistical calculations of the outcomes. The two teachers, who evaluated the scoring rubric in the preparatory stage, worked as independent raters.

#### **3.6.3.1 Preparatory stage**

Some preparatory work preceded scoring and coding the compositions. All 390 compositions were typed and saved as Word files to standardize the essay format and

avoid handwriting bias; any information about identification code, draft condition (Draft 1/2), and group condition (treatment/comparison) was removed.

The two independent raters received training in using the scoring rubric (Appendix A) and the error coding scheme (Appendix F). In the first half of the training session, I explained the adapted scoring rubric and the procedures for scoring; I then invited the two raters to mark 10 compositions representing different levels of performance independently. They discussed any discrepant scores by referring to the scoring criteria as well as my reference scores. The two raters then scored 30 randomly selected compositions (around 8% of the total compositions) to determine interrater reliability. The interrater reliability for content, organization, vocabulary, grammar, and total scores was .722, .288, .827, .674 and .777, respectively. As main differences were in organization and grammar scores, the two raters and I discussed discrepant scores in the two categories to solve the disagreements. The two raters then rated another 30 drafts independently. The interrater reliability for content, organization, vocabulary, grammar, and total score reached .788, .864, .794, .812 and .824, respectively.

The second half of the training session focused on coding errors. Similar to the training on scoring the compositions, I explained the error coding scheme adapted from Geng (2017) before inviting the two raters to code 30 randomly selected compositions for errors independently. The interrater reliability for calculating errors was .764. The two raters discussed and solved their disagreements before continuing with another 30 compositions. This time the interrater reliability improved to .838.

Since the interrater reliability for the writing scores (overall score as well as scores in the four specific dimensions) and for calculating errors reached the threshold level of .70 (Hamp-Lyons, 1990), the scoring and coding stage proceeded.

### 3.6.3.2 *Scoring and coding the compositions*

The 390 compositions were divided equally between the two raters, who scored the compositions using the adapted Jacobs et al.'s (1981) scoring rubric and coded errors according to the error coding scheme. To ensure an equal effect of training, I requested the two raters to carry out the scoring task and the coding task simultaneously.

#### 3.6.3.2.1 Assessing the overall quality, content quality and organization quality

The adapted Jacobs et al.'s (1981) scoring rubric was adopted in this study to evaluate the overall quality, content quality and organization quality of a composition. The rubric uses “a weighted scoring scheme on a percentile scale” (Teng & Zhang, 2016, p.129), measuring the five dimensions of writing performance: content (30%), organization (20%), vocabulary (20%), grammar (25%), and mechanics (5%). For each dimension, there are four rating scales, each with specific criteria.

The decision to consult Jacobs et al.'s (1981) scoring rubric was made based on four reasons. Firstly, this analytic scoring rubric could “provide more detailed information about a test taker’s performance in different aspects of writing” (Weigle, 2002, p. 114). Secondly, since its introduction, this scoring rubric has been widely and successfully used to evaluate compositions written by L2 learners (e.g., Huang & Zhang, 2020; Sasaki, Mizumoto, & Murakami, 2018). Thirdly, the weighting of different dimensions matches the foci in writing instruction in EFL contexts. The percentage taken by vocabulary, grammar and mechanics together (50%) is consistent with the emphasis on language use, accuracy in particular, in EFL writing instruction and evaluation. Fourthly, this scoring rubric is user-friendly in that it includes explicit explanations for each criterion as well as 18 writing samples as training materials.



### 3.6.3.2.2 Assessing language use quality

In the present study, I employed some CAF indexes to assess the quality of language use. Specifically, mean length of T-units (MLT) is used to measure syntactic complexity; mean segmental type-token ratio per 50 words (MSTTR\_50) is used to measure lexical complexity; the number of errors per 100 words (EP100W) is used to measure accuracy; and the number of words per text (WPT) is used to measure fluency.

Complexity is defined as “the extent to which the language produced in performing a task is elaborate and various” (Ellis, 2003, p. 340); it can be further categorized as syntactic complexity and lexical complexity (Vyatkina, 2012). In this study, MLT was chosen because it was a global measure and some researchers could consider it a satisfactory predictor of writing quality (Casal & Lee, 2019; Ortega, 2015; Wolfe-Quintero et al., 1998). To calculate MLT, a T-unit should first be defined. T-unit is defined as “one main clause plus whatever subordinate clauses happen to be attached or embedded with it” (Hunt, 1966, p. 737). MLT is calculated by dividing the total number of words by the number of T-units in a composition. The higher the number, the greater the syntactic complexity of a composition.

Lexical complexity used to be measured by the type-token ratio, obtained by dividing the number of different words in a text by the total number of its words. This measure, however, has been severely criticised because of its “loss of discriminatory power over sample size” (deBoer, 2014, p. 140). Mean segmental TTR (MSTTR), as an alternative measure, has been introduced. To calculate MSTTR, the text is first divided into segments (for example, 40 or 50 words each), type-token ratio of each segment is then computed, and finally the mean score for all the segments is calculated. In this study, the mean segmental type token ratio per 50 words (MSTTR\_50) was used.

Accuracy refers to the “target-like use [of language], taking into account both the contexts and uses of the structure in question” (Wolfe-Quintero et al., 1998, p. 33). To calculate the number of EP100W, the total number of errors is divided by the total number of words in a composition, and then multiplied by 100.

Fluency is defined as “the learner’s capacity to mobilize an interlanguage system to communicate meanings in real time” (Skehan, 1996, p. 46), and in the mode of writing, fluency is measured by “the number, length, and rate of production units” (Wolfe-Quintero et al., 1998, p. 14). In this study, WPT is obtained by counting the number of words in a composition.

In coding the collected compositions for the CAF measures, the data for MLT, MSTTR\_50 and WPT were automatically generated by the Web-based L2 Syntactic Complexity Analyzer (<http://aihaiyang.com/software/l2sca/>), whereas the data for EP100W were manually calculated.

### 3.6.3.3 *Statistical analysis*

This section provides an overview of the statistical methods used in this study. In Chapter 4 which reports the result for this part of study, relevant statistical analyses will be reviewed briefly.

Numeric data from scoring and coding the compositions were submitted to a series of statistical analyses using the IBM Statistical Package for Social Sciences (SPSS) 22. Specifically, the numeric data were transcribed into two datasets, one containing the writing scores and the CAF indexes for the treatment group and comparison group participants’ first and second drafts of Writing Task 6, and the other containing the

writing scores and the CAF indexes for their compositions in the pre-test, post-test, and delayed post-test.

Before each statistical test was run, the relevant data were examined for underlying assumptions. The Shapiro-Wilk's test, visual inspection of histograms, normal Q-Q, box plots, and skewness and kurtosis were used to check for the normality of distribution; Levene's test was used to check for homogeneity of variance, and Mauchly's test was used to check for sphericity. The significance level was set at .05. For multiple comparisons, Bonferroni corrections were applied to avoid Type I errors. Effect size, to ascertain the practical significance of the results, was calculated. Partial eta squared ( $\eta^2$ ) was reported for mixed ANOVA and ANCOVA; and Cohen's *d* was calculated for independent and paired-samples t-tests. The effect size was interpreted according to Cohen (1988):  $\eta^2$  values of .01, .06, .14 and *d* values of .20, .50, and .80 were considered small, medium, and large, respectively.

To respond to the research question concerning the effects of peer feedback on text revisions, ANCOVA and paired-sample t-test were employed to examine between-group differences and within-group differences, respectively. ANCOVA was used to control for pre-existing differences between the two groups. Since the data were collected from Writing Task 6, after the treatment group had participated in five writing tasks with peer feedback activities, it was possible that the two groups differed with some measures in their first drafts. To establish within-in group differences, a paired-sample t-test was applied to examine changes within each group between their first and second drafts.

A mixed ANOVA was first applied to each measure to examine main effects and interaction effect to establish the effect of peer feedback on writing performance. To

further explore within-group differences and between-groups differences, one-way repeated measure ANOVA and independent t-test were deployed respectively.

### **3.7 Part Two: Processing of peer feedback in revision**

Part Two was a multiple case study which explored how the case study participants made use of peer feedback when writing their second drafts.

Based on data collected from various sources in Writing Task 6, this part of study addressed the third research question and its sub-questions.

RQ 3 How do Chinese EFL learners process peer feedback when writing their second drafts?

*What cognitive operations do Chinese EFL learners employ when processing peer feedback points?*

*What approaches do Chinese EFL learners take when processing peer feedback points?*

Case study, characterized by boundedness and in-depth analysis, is appropriate for investigating how individuals work in a particular context or situation with a focus on “a contemporary set of events over which the investigator have little or no control” (Yin, 1994, p. 9). To conduct a case study, researchers commonly focus on two to six multiple cases and use multiple instruments and data sources to generate thick descriptions (Duff, 2020).

The present study took a multiple case approach by including six non-English major students to investigate how EFL learners made use of peer feedback when making revisions to their writing. The multiple case approach was used with the aim of capturing

individual variations and minimizing the risk of attrition of participants (Dörnyei, 2007; Merriam & Tisdell, 2015).

### 3.7.1 **Participants**

Six participants, who were in the treatment group receiving the intervention, were recruited for the multiple case study through convenience sampling (see 3.6.1). There were two criteria for recruitment: demonstrating the ability in performing think-alouds and having sufficient time to attend the think-aloud training and the two think-aloud sessions. No consideration was given to gender, major, or writing proficiency.

### 3.7.2 **Data collection**

Qualitative data were obtained from the six case study participants in Writing Task 6. Data included their think-aloud protocols, transcripts of paired discussions, as well as a set of written documents, including their two drafts for Writing Task 6 and the accompanying PF forms.

#### 3.7.2.1 *Think-aloud protocols*

Think-aloud was employed as a method for data collection to gain insights into participants' processing of peer feedback during revision. As an introspective method of "gathering information about human information processing" (L. J. Zhang & Zhang, 2020, p. 302), this method requires participants to report "what is going through their minds as they are solving a problem or performing a task" (Gass & Mackey, 2000, p. 13). In comparison with stimulated recall interviews, which require participants to recall an activity after they completed a task, and thus having the drawbacks of the forgetting effect and post hoc rationalizations (Sanchez & Grimshaw, 2020), think-aloud protocols provide a rich amount of real-time data underlying the participants' task performance

(Charter, 2003; Gass & Mackey, 2000; L. J. Zhang & Zhang, 2020). With a strong theoretical basis in working memory and information processing, think-aloud has been widely employed in the fields of psychology, cognitive science, and applied linguistics (Ericsson & Simon, 1993). In L1 and L2 writing research, think-aloud is of great value, as it reveals participants' decisions, strategies, and perceptions in the writing process (Huang, 2020; Sasaki, 2000; Suzuki, 2008).

Two issues, reactivity and veridicality, however, should be noted in the use of think aloud protocols. Reactivity refers to possibility of changes in observed behaviors or outcomes caused by performing the dual task of writing and verbalizing thoughts at the same time; veridicality concerns the possible discrepancies between the participants' actual cognitive process and their reported ones (C. Yang, Zhang, & Parr, 2020; L. J. Zhang & Zhang, 2020). To minimize the potential threats related to these two issues, a group training session as well as individual training sessions were provided for the six case study participants in Week 11, two weeks prior to the collection of the think-aloud data.

During the group training session, the rationale and purpose for using think-alouds were first explained, as recommended by L. J. Zhang and Zhang (2020). Also, specific instructions on how to conduct think-alouds were provided. The following points were highlighted as advocated by Perkins (1981).

- (1) Say whatever's on your mind. Don't hold back hunches, guesses, wild ideas, images, intentions.
- (2) Speak as continuously as possible. Say something at least once every five seconds, even if only, "I am drawing a blank".

(3) Speak audibly. Watch out for your voice dropping as you become involved.

(4) Speak as telegraphically as you please. Don't worry about complete sentences and eloquence.

(5) Don't overexplain or justify. Analyze no more than you would normally.

(6) Don't elaborate past events. Get into the pattern of saying what you are thinking now, not only thinking for a while and then describing your thoughts.

(p.33)

In addition to the six points, the case study participants were told to articulate their thoughts when writing their second drafts, in Chinese, English, or a mixture of both.

A video demonstrating how a writer made revisions to a paragraph with feedback while thinking aloud followed. After watching the video, the six case study participants were given opportunities to ask any questions regarding think-aloud techniques and other related matters.

As well as the group training session, I scheduled an individual practice session with each participant, in which the participant made revisions to his/her first draft of Writing Task 5 while thinking aloud. The practice sessions resembled the actual think-aloud sessions in every aspect to familiarize the case study participants with the dual tasks of thinking aloud and writing their second drafts. The six participants took an average of 70 minutes to finish their second drafts. The audio recordings indicate that all participants understood the techniques, as they produced clear verbalizations and reported no difficulty in doing such a task. At the end of each practice session, the participant was reminded of specific problems and provided with suggestions.

The collection of think-aloud data occurred in the first week of Writing Task 6 (Week 13). All the six sessions were carried out in a quiet classroom individually, during which I sat in the back corner of the classroom to avoid distracting the participants. I only intervened when the participants paused for over five seconds or lowered their voice too much. When such situations occurred, I reminded them by saying “Please keep talking” and “Please raise your voice” respectively. All six case study participants mainly used Chinese when thinking aloud their revision process. Approximately nine hours of think-aloud data were audiotaped with their permission.

### 3.7.2.2 *Documents from Writing Task 6*

Documents serve as important data sources for case studies as they contain exact and factual details of an event, and can be repeatedly examined (Yin, 2015). In this part of study, documents obtained from Writing Task 6 included the case study participants’ first drafts with feedback, the accompanying PF forms, and their second drafts. These documents provided data for analysis of feedback and therefore helped identify revision episodes triggered by peer feedback.

### 3.7.2.3 *Paired discussion transcripts*

Paired discussions, in which case study participants discussed written peer feedback with their peers, were audio-recorded to provide data for feedback and revision analysis. Since the six case study participants were randomly paired with other students who were not included in the case study, only the discussions on the case study participants’ first drafts were recorded. Approximately two hours of paired discussions were audiotaped with permission from the six case study participants and their peers.



#### 3.7.2.4 *Transcription conventions*

To ensure that the results were clearly displayed, I established my own transcription conventions as follows:

- Block letters: Discourse from paired discussions and think-aloud protocols spoken in Chinese; feedback in the PF form or on the draft, written in Chinese
- Bold block letters: Discourse from paired discussions and think-aloud protocols, spoken in English; feedback in the PF form or on the draft, written in English
- Bold italics in angle brackets: cognitive operations for specific excerpts from think-aloud protocols
- Square brackets: the researcher's annotations (describing a situation, showing the reference to a written text, and translating the Chinese expressions into English)
- Round brackets: the source of the information

There were two further points to note about the transcription: When the case study participants verbalized a misspelled word or a grammatical error, it was recorded as originally appeared in their speech or writing; the names for all case study participants were pseudonyms, with their partners in discussions referred to as Peer.

### 3.7.3 **Data analysis**

#### 3.7.3.1 *Analysis of peer feedback received*

The feedback dataset, containing 100 feedback points, was established before the analysis started. It consisted of all feedback points that the six case study participants received on their first drafts of Writing Task 6, including the written feedback points on

their first drafts and on the accompanying PF forms and the oral feedback points exchanged in paired discussions. Feedback points addressing the same problem were combined but compliments triggering no revisions were excluded.

Table 3.7 Coding scheme for feedback points

---

<b>Feedback focus</b>
<b><u>Content:</u></b> Feedback addressing problems of information relevance, adequacy, and sufficiency
<b><u>Organization:</u></b> Feedback addressing problems of global text structure and local text structure
<b><u>Language use:</u></b> Feedback addressing problems in grammar (e.g., mistakes in subject-verb agreement and misuse of tense), vocabulary (e.g., proper use of a word), mechanics (e.g., misspelling or inappropriate punctuation)

---

<b>Feedback validity</b>
<b><u>Valid feedback:</u></b> All information within a feedback point is correct
<b><u>Invalid feedback:</u></b> Some information within a feedback point is inaccurate or inappropriate.

---

All feedback points were classified according to two criteria: focus and validity. The first classification represented the criteria used to evaluate the quality of compositions and included feedback on content, organization and language use. Each feedback point was also classified as valid or invalid feedback. The coding scheme for feedback points is shown in Table 3.7.

To illustrate the coding of feedback points, I took the following feedback points as an example.

(Shen, received feedback on PF form)

“**Literature**” cannot be the subject [in the sentence “On the other hand, literature improves one’s ability to persist”]. You can use “**reading literature**”.

(Discussion between Shen and his peer)

**Peer:** **Literature** cannot improve your abilities but **reading literature** can.

**Shen:** Yes, you are right. Literature cannot improve abilities, but reading literature can.

Firstly, the two pieces of feedback, one from Shen’s PF from and the other one from the discussion were combined and counted as one feedback point in the dataset, as they address the same problem. In terms of focus, this feedback point focuses on language, to be specific, word choice. In terms of validity, this feedback point was valid, because both the problem and the solution were correctly identified in the written feedback; the explanation offered by the reviewer in the discussion was also correct.

When I finished coding all the feedback points, I shared the transcripts of feedback points with one rater and invited her to code them. We agreed upon all feedback points in terms of focus and 96% of the feedback points in terms of validity. All disagreements were solved through discussions.

### 3.7.3.2 *Analysis of cognitive operations in PF-triggered revision episodes*

In this study, “revision episode” is defined as any part of think-aloud protocols in which a case study participant addresses a problem in his/her composition, similar to “language related episode” defined by Swain and Lapkin (1998) and “negotiation episode” defined by Suzuki (2008). Each revision episode consists of cognitive operations, being referred

to as actions writers undertook to make the change(s). While self-regulation theories and second language learning strategy literature used strategies (e.g., Oxford, 2016; O’Malley & Chamot, 1990), I decided to use “operation” rather than “strategy” in this study, to be consist with the literature on revision models (see Alamargot & Chanquoy, 2001).

Of all the revision episodes, this study, with a focus on peer feedback, only examined PF-triggered revision episodes in which a participant explicitly mentioned part(s)/whole parts of a received feedback point. The remaining were categorized as other revision episodes and excluded from this study, because I cannot decide what triggered these revisions based on the think-aloud data. Content analysis, a method which “code text into categories and then count the frequencies of occurrences within each category” (Ahuvia, 2001, p. 139), was used to analyze cognitive operations in PF-triggered revision episodes.

First, I manually segmented each think aloud protocol into revision episodes and identified PF-triggered revision episodes based upon feedback data in 3.7.3.1. The initial analysis yielded a total of 102 revision episodes, with 82 PF-triggered revision episodes. The 82 PF-triggered revision episodes were numbered individually for reporting of the findings.

For the coding of PF-triggered revision episodes for cognitive operations that followed, I adopted a deductive approach based on the coding schemes in two teacher feedback studies (Han, 2016; Rajoo, 2009). I started by repeatedly reading all PF-triggered revision episodes from the six case study participants, and then coded PF-triggered episodes from two case study participants, Shen and Xu. When coding, I modified the coding schemes of Rajoo’s (2009) and Han’s (2016) studies by removing some

cognitive operations (e.g., predicting and memorizing), redefining some cognitive operations (e.g., evaluating), as well as adding some cognitive operations (e.g., translating and resourcing), to describe the current data better. My coding scheme and coding of the case study participants' cognitive operations were shared with one rater. Based on her feedback, I modified my coding and finalized my coding scheme, as shown in Table 3.8., with definition and example for each cognitive operation.

Of the nine cognitive operations, *retrieving L2 knowledge* and *monitoring* were borrowed from Han (2016) and *justifying* and *generating changes/no change* were borrowed from Rajoo (2009). *Referring to feedback* was adapted from Rajoo (2009) by combining the three cognitive operations, *reading feedback (exact words)*, *rereading or reviewing parts of whole/parts of feedback*, and *referring to feedback (not exact words)*. Firstly, the distinction between reading (the exact words) and referring (not the exact words) were unnecessarily refined. Secondly, because my study focused on how frequently that the case study participants used each cognitive operation, I did not need to distinguish between reading and rereading. When students reread or reviewed part(s)/whole of feedback that they received, they were coded as referring to feedback again, just as some cognitive operations were used more than once in a PF-triggered revision episode. For a similar reason, *referring to text* was adapted from Rajoo (2009) by combining the three cognitive operations, *reading parts of essay/sentences/section*, *rereading or reviewing parts of essay/sentences/section* and *referring to parts of essay/sentences/section (not exact words)*. *Evaluating* in this study was defined based on data from this study, different from that in Han (2016) in which students evaluated their revising process or gravity of their errors after they made all revisions to their text. The two cognitive operations, *translating* and *resourcing*, not present in Han (2016) or Rajoo (2009), were defined according to O'Malley and Chamot (1990).

Table 3.8 Coding scheme for cognitive operations in PF-triggered revision episodes

Cognitive operations	Definition	Example
<b>Referring to feedback</b>	A participant is reading or paraphrasing a received feedback point or putting a feedback point back to a corresponding text segment.	In this sentence, “ <b>In my opinion, the mother’s view is considerable and wise</b> ”, “ <b>in my opinion</b> ” is too common. He [Her peer] suggested that I use a complex or advanced expression. (Liu, PF-triggered revision episode 2)
<b>Referring to text</b>	A participant is reading or paraphrasing part of a text.	I wrote [in his first draft], “ <b>When we are dropped in the book we are reading, we can feel what the character feels and identify what’s good or bad for us to accept</b> ”. (Yang, PF-triggered revision episode 4)
<b>Resourcing</b>	A participant is consulting reference materials (e.g., dictionaries and grammatical books).	[Using the Chinese-English translation of an online dictionary to find the English equivalent for “割裂的”] “Split”, there are “ <b>splitted</b> ” [ <i>sic</i> ] and “ <b>intersected</b> ”. “ <b>Intersect</b> ” means to cut and to divide [ <i>sic</i> ]. There is another word, “ <b>divisional</b> ”. (Sun, PF-triggered revision episode 15).
<b>Translating</b>	A participant is using L1 (Chinese) as a base for understanding or producing a text segment.	I will write “为什么阅读被放在了一个这么重要的位置上”, “为什么阅读放在了这么一个重要的位置上”. (Liu, PF-triggered revision episode 10)

<b>Retrieving L2 knowledge</b>	A participant is recalling previous knowledge about L2 and L2 writing.	How could I express “复杂的”[meaning complicated in Chinese]? I can say “ <b>complex social background</b> ”. “ <b>Complicated</b> ” also had a similar meaning. (Yang, PF-triggered revision episode 3)
<b>Monitoring</b>	A participant is tracking his/her attention to a text segment when the revision is taking place.	I have used “have a great influence”. I need to use another word with a similar meaning. (Zhang, PF-triggered revision episode 5)
<b>Justifying</b>	A participant is providing reasons for a received feedback point or for changes to be made.	Because the following paragraph talked about knowledge, “ <b>train a person to think</b> [ <i>sic</i> ]” could not summarize what I wrote. (Shen, PF-triggered revision episode 3)
<b>Generating changes/no change</b>	A participant decides to implement change(s) or retain what he/she has written.	I will delete this sentence [“Only because phones are quickly to search it can’t be the reason we keep using it”]. (Wu, PF-triggered revision episode 4)
<b>Evaluating</b>	A participant is checking a revision change against a standard (received feedback/their own judgement about its appropriateness).	[After deleting “on the one hand” in the sentence “On one hand, by providing systematically knowledge in some fields, literature helps to develop the comprehensive and critical thinking in one’s mind.”] The new sentence was coherent with its previous sentence. (Shen, PF-triggered revision episode 4)

To illustrate the coding procedure, I presented a section of Shen's think-aloud transcript with the categorizations of revision episodes and cognitive operations for PF-triggered revision episodes as an example (see Appendix G). When I finished coding all the revision episodes, I shared with each case study participant my coding scheme and coding of his/her PF-triggered revision episodes for member checking.

### **3.8 Part Three: Perceptions of peer feedback**

Part three of the study included a questionnaire involving the 35 treatment group participants and interviews with the six case study participants. Specifically, it addressed the following research question:

RQ 4 How do Chinese EFL learners perceive effectiveness of the peer feedback intervention?

*How do Chinese EFL learners perceive their changes in writing performance?*

*How do Chinese EFL learners perceive the usefulness of specific components in the peer feedback intervention?*

Questionnaires, as one of the most popular instruments for social science research, have been used to collect data about various topics in L2 writing research, such as students' views and experience about writing (e.g., Ferris & Eckstein, 2020), their attitudes and opinions about different feedback practices (e.g., Rajoo, 2009), and their strategies in writing (e.g., Teng & Zhang, 2020). Compared with other data collection methods, questionnaires are advantageous in several ways: Questionnaires are efficient in terms of researchers' time and effort because they can be distributed to a group of people at the same time; questionnaires can be various in formats by including Likert scales, semantic differential scales, numerical rating scales, or open-questions, to suit different



research objectives; questionnaires do not involve interviewer bias, which may threaten reliability of the findings; data collected from well-constructed questionnaires can be processed fast using various software programs (Iwaniec, 2020).

For this part of study, open-ended questions were included in the questionnaire in that they permit greater freedom of expression and elicit more abundant data than closed items (Dörnyei & Taguchi, 2009). There are, however, some potential problems with incorporating open-ended questionnaires: The response time may limit the participants' range of answers; some of their responses may be superficial or even irrelevant. To mitigate such problems, this study included only nine open-ended questions with clear directions and provided the participants with adequate time to answer them. Also, semi-structured interviews, which allow interviewers to guide the interviewees to elaborate on their responses (Dörnyei, 2007), were employed in this study to bring about in-depth perceptions not revealed through the questionnaire.

### **3.8.1 Data collection**

The Perceptions of Peer Feedback Questionnaire and six interviews were used to collect data for this part of study. The questionnaire was prepared in Chinese (Appendix H). Each item included a Likert scale and some space eliciting the respondents' explanations for their options on the Likert scale. An example of questionnaire items is as follows:

By participating in peer feedback activities all through the semester, I have improved in the content of my composition.

1	2	3	4	5
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

Questionnaires were distributed to the 41 treatment group participants during class time in Week 16. When administering the questionnaires, the teacher reminded the participants to make a choice on the five-point Likert scale and provide a detailed explanation for their choice.

Questionnaires of all the 41 treatment group participants were returned to the researcher. Because the questionnaires in which the participants skipped four or more questionnaire items (making no choice on the Likert scale and/or providing no explanation for their choice) were excluded, a total of 35 questionnaires were subject to data analysis.

Semi-structured interviews were conducted with each case study participant individually in Week 16. Approximately four hours of interviews were audiotaped with their permission. The interview questions (see Appendix I), resembling those in the Perceptions of Peer Feedback Questionnaire, elicited their opinions on the impact of peer feedback and on the usefulness of specific activities of the intervention. To facilitate the case study participants' responses, I prepared all their written documents for the six writing tasks. All interviews took place in a quiet café, and were conducted in Mandarin, the participants' native language.

### 3.8.2 Data analysis

The participants' choices on the Likert scale were analyzed quantitatively. However, as this study focused more on the treatment group participants' explanations, the five levels on the Likert scale (1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, and 5=strongly agree) were reorganized into three, "disagree" (including "strongly disagree" and "disagree" responses on the Likert scale), "neither disagree nor agree", and "agree" (including "strongly agree" and "agree" responses on the Likert scale). For each item, I tallied the treatment group participants' choices for each of the three levels and calculated it as a percentage.

The explanations that the treatment group participants provided for each questionnaire item were given priority in the data analysis using content analysis; each item was analyzed individually taking an inductive approach. For each questionnaire item, I first read through all participants' explanations several times to identify topics. When the topics were refined and established, I coded their explanations into categories and counted frequencies for each category. Because there were instances in which some participants skipped some questions and other instances in which their explanations touched upon multiple topics and were therefore segmented and put into different categories, the number of responses for each question was not 35, with some fewer than 35, and some more than 35.

To ensure the trustworthiness in coding, I shared my topics for each question with one rater and invited her to code the treatment group participants' responses. We agreed upon 91.1% of the responses and resolved our disagreements through discussions.

All interview transcripts were first segmented into episodes according to the research questionnaire items. I then reviewed them a few times to identify some informative and

representative ones. These episodes were presented to complement those selected from the questionnaires.

### **3.9 Ethical considerations**

#### **3.9.1 Informed consent**

Prior to the field work, I sought consent from the Dean, the teacher and student participants. They were informed that the research project was a part of my PhD research. To indicate they fully comprehended the purpose and procedures of the research project, as stated in the Participant Information Sheet (PIS), they signed consent forms (CF). All PISs and CFs are in Appendix J.

#### **3.9.2 Right to withdraw from participation**

Participants were also informed that their participation was completely voluntary, and they were entitled to withdraw either themselves or data provided by them without giving any reasons within three weeks after the data collection was completed. They were also assured that their withdrawal would neither disadvantage them nor affect their grades for the College English course.

#### **3.9.3 Confidentiality**

Confidentiality was guaranteed during the whole research process. Each participant was assigned a unique identification code and was required to use it consistently for all the written documents for the study. Participants were also assured of anonymity in reporting of research findings. If any information provided by the participants is reported or published, participant numbers or pseudonyms will be used to protect their identity; no identifying information or data collected for this research project will be disclosed to a third party.

## Chapter 4 Effects of Peer Feedback Intervention

### 4.1 Chapter overview

This chapter reports findings regarding the effects of peer feedback on text revisions and writing performance, responding to the following two research questions and their sub-questions.

RQ 1 How does peer feedback influence Chinese EFL learners' text revisions?

*How does peer feedback influence the overall quality, content quality and organization quality of Chinese EFL learners' text revisions?*

*How does peer feedback influence the accuracy, syntactic complexity, lexical complexity, and fluency of Chinese EFL learners' text revisions?*

RQ 2 How does peer feedback influence Chinese EFL learners' writing performance?

*How does peer feedback influence Chinese EFL learners' writing performance in terms of the overall quality, content quality and organization quality?*

*How does peer feedback influence Chinese EFL learners' writing performance in terms of complexity, accuracy, and fluency?*

Comparisons of the two drafts of Writing Task 6 written by the treatment group and comparison group participants are presented to show the effects of peer feedback on text revisions. The compositions written by the treatment group and the comparison group participants in the pre-test (Week 1), the post-test (Week 16), and the delayed post-test (Week 25) were compared to show the effects of peer feedback on writing performance.

There are four sections in this chapter. Following the chapter overview, the two sections report the effects of peer feedback on text revisions and on writing performance respectively. The final section summarizes the findings in this chapter.

## **4.2 Effects of peer feedback on text revisions**

This section reports findings of changes in writing scores (overall score, content score, and organization score), and CAF indexes (EP100W for accuracy, MLT for syntactic complexity, MSTTR\_50 for lexical complexity and WPT for fluency) between drafts. Paired-samples t-tests were performed to examine within-group differences, and ANCOVAs were applied to examine between-group differences.

### **4.2.1 Comparison of baseline conditions of the two groups**

The participants' first drafts of Writing Task 6 served as the baseline data for examining the effects of peer feedback on text revisions. After checking the assumptions of normal distribution and homogeneity, one-way ANOVAs were conducted to compare the treatment group and the comparison group participants' first drafts in terms of writing scores and CAF indexes. As seen in Table 4.1, the treatment group significantly outperformed the comparison group in overall score ( $F(1, 76) = 9.16, p = .003$ ) and in organization score ( $F(1, 76) = 4.28, p = .042$ ) of their first drafts. For written content of Draft 1, although the treatment group had a higher score than the comparison group, the difference was not statistically significant ( $F(1, 76) = 3.97, p = .05$ ).

Table 4.1 Descriptive statistics and one-way ANOVA analysis of scores for Draft 1 of Writing Task 6

Scores	Group	N	M	SD	F	df	p
Overall	Treatment	41	73.10	5.22	9.16	76	.003*
	Comparison	37	69.54	5.14			
Content	Treatment	41	21.66	2.42	3.97	76	.050
	Comparison	37	20.65	2.02			
Organization	Treatment	41	15.05	1.34	4.28	76	.042*
	Comparison	37	14.41	1.40			

Notes. \*  $p < .05$

Table 4.2 shows descriptive statistics and one-way ANOVA analysis of CAF indexes for Draft 1 of Writing Task 6. For their first drafts, the treatment group significantly outperformed the comparison group in accuracy ( $F(1, 76) = 15.17, p < .001$ ), but the two groups performed similarly in syntactic complexity ( $F(1, 76) = 1.43, p = .236$ ), lexical complexity ( $F(1, 76) = .13, p = .725$ ), and fluency ( $F(1, 76) = .17, p = .682$ ).

The differences in some scores and CAF indexes between the two groups are in accordance with my expectations. As the treatment group participants wrote their first drafts of Writing Task 6 after they participated in five writing tasks with peer feedback activities, it is understandable that they significantly outperformed the comparison group participants in some measures.

Table 4.2 Descriptive statistics and one-way ANOVA analysis of CAF indexes for Draft 1 of Writing Task 6

CAF indexes	Group	N	M	SD	F	df	p
<b>Accuracy</b> EP100W	Treatment	41	6.34	1.91	15.17	76	.000**
	Comparison	37	8.18	2.28			
<b>Syntactic complexity</b> MLT	Treatment	41	15.96	2.66	1.43	76	.236
	Comparison	37	16.73	3.10			
<b>Lexical complexity</b> MSTTR_50	Treatment	41	.78	.03	.13	76	.725
	Comparison	37	.78	.04			
<b>Fluency</b> WPT	Treatment	41	252.93	38.98	.17	76	.682
	Comparison	37	248.81	49.29			

Notes. EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text; \*\*  $p < .001$

#### 4.2.2 Changes in writing scores between drafts

Table 4.3 displays the descriptive results of overall scores, content scores, and organization scores for the two drafts of Writing Task 6. For each writing score for Draft 1 and Draft 2, the mean for the treatment group was higher than that for the comparison group. Both groups increased their writing scores from Draft 1 to Draft 2.



Table 4.3 Group means and standard deviations of the overall scores for Draft 1 and Draft 2 of Writing Task 6

Scores	Group	N	Draft 1		Draft 2	
			M	SD	M	SD
<b>Overall</b>	Treatment	41	73.10	5.22	76.78	4.74
	Comparison	37	69.54	5.14	72.19	5.35
<b>Content</b>	Treatment	41	21.66	2.42	22.73	2.66
	Comparison	37	20.65	2.02	21.11	2.08
<b>Organization</b>	Treatment	41	15.05	1.34	15.76	1.04
	Comparison	37	14.41	1.40	14.73	1.28

*Notes. D1 = Draft 1 of Writing Task 6; D2 = Draft 2 of Writing Task 6*

A paired-samples t-test was run to examine the within-group differences in overall score, content score, and organization score between drafts. As can be seen from Table 4.4, each group increased significantly in overall score between drafts ( $p < .001$  for the treatment group, and  $p < .001$  for the comparison group), in content score between drafts ( $p < .001$  for the treatment group, and  $p = .045$  for the comparison group). The organization score of each group also increased significantly between drafts ( $p < .001$  for the treatment group and  $p < .001$  for the comparison group). For each score, the effect size  $d$  for the treatment group was larger than that for the comparison group (1.16 vs. 0.78 for overall score, 0.90 vs. 0.34 for content score, and 0.75 vs. 0.36 for organization score).

Table 4.4 Within-group differences in scores between drafts of Writing Task 6

Measure	Group	Mean	SD	95% CI		t	df	p
				LL	UL			
<b>Overall</b>	Treatment	-3.68	3.15	-4.68	-2.69	-7.49	40	.000**
	Comparison	-2.65	3.43	-3.79	-1.51	-4.70	36	.000**
<b>Content</b>	Treatment	-1.07	1.20	-1.45	-.70	-5.77	40	.000**
	Comparison	-.46	1.35	-.91	-.01	-2.08	36	.045*
<b>Organization</b>	Treatment	-.71	.96	-1.01	-.41	-4.74	40	.000**
	Comparison	-.32	.88	-3.79	-1.51	-4.70	36	.000**

Notes. \*  $p < .05$ ; \*\*  $p < .001$

A series of ANCOVAs were performed to examine between-group differences in writing scores for Draft 2 of Writing Task 6. Basic assumptions of ANOCVA were examined first in terms of homogeneity of variance and homogeneity of regression. The samples were found to have homogenous variances for overall, content, and organization scores; no violation was found in terms of the homogeneity of regression for overall, content, or organization scores.

ANCOVA results indicate significant differences in overall score ( $F(1, 75) = 6.23, p = .015, \eta^2 = 0.08$ ), in content score ( $F(1, 75) = 5.68, p = .02, \eta^2 = 0.07$ ), and in organization score ( $F(1, 75) = 11.62, p = .001, \eta^2 = 0.13$ ) for Draft 2 between the treatment group and the comparison group. The treatment group significantly outperformed the comparison group, with medium effect sizes, in overall score, content score, and organization scores for Draft 2, when differences in corresponding measures for Draft 1 were controlled.

To summarize, both groups made significant improvements in overall, content, and organization qualities between the two drafts of Writing Task 6. In addition, the treatment group made greater gains than the comparison group in overall, content and organization scores.

#### **4.2.3 Changes in CAF indexes between drafts**

Table 4.5 displays the descriptive results of the four CAF indexes for the two drafts of Writing Task 6. In terms of accuracy, both groups decreased in the number of errors between drafts; the treatment group made fewer errors than the comparison group in Draft 1 and in Draft 2. The syntactic complexity of the treatment group's drafts improved more than the comparison group, but the lexical complexity of the two drafts of the two groups was similar. Both groups increased in fluency between drafts; the treatment group produced longer texts for Draft 1 and Draft 2 than the comparison group.

Table 4.5 Group means and standard deviations of CAF indexes for Draft 1 and Draft 2 of Writing Task 6

CAF Indexes	Group	N	Draft 1		Draft 2	
			M	SD	M	SD
<b>Accuracy</b> EP100W	Treatment	41	6.34	1.91	5.22	1.63
	Comparison	37	8.18	2.28	6.98	2.35
<b>Syntactic complexity</b> MLT	Treatment	41	15.96	2.66	16.60	3.06
	Comparison	37	16.73	3.10	16.74	2.90
<b>Lexical complexity</b> MSTTR_50	Treatment	41	.78	.03	.78	.04
	Comparison	37	.77	.03	.78	.03
<b>Fluency</b> WPT	Treatment	41	252.93	38.98	261.17	39.29
	Comparison	37	248.81	49.29	250.30	39.84

Notes. EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text

The paired-samples t-test result for each of the four CAF indexes is shown in Table 4.6. As seen in the table, both groups decreased significantly in the number of errors from Draft 1 to Draft 2 with large effect sizes ( $p < .001$ ,  $d = 0.86$  for the treatment group, and  $p < .001$ ,  $d = 1.07$  for the comparison group), that is, both groups improved in accuracy between drafts and the comparison group made greater improvements than the treatment group. The treatment group significantly increased in MLT from Draft 1 to Draft 2, with a small effect size ( $p = .015$ ,  $d = .40$ ), indicating increased syntactic complexity; there was no change in MLT between drafts for the comparison group ( $p = .386$ ), however. Neither the treatment group nor the comparison group made significant changes between drafts in lexical complexity ( $p = .512$  for the treatment group and  $p = .947$  for the comparison group). With respect to fluency, there was no significant change in text

length between drafts either for the treatment group ( $p = .085$ ) or the comparison group ( $p = .715$ ).

Table 4.6 Within-group differences in CAF indexes between drafts of Writing Task 6

CAF indexes	Group	Mean	SD	95% CI		t	df	p
				LL	UL			
<b>Accuracy</b> EP100W	Intervention	1.12	1.30	.71	1.53	5.52	40	.000**
	Comparison	1.21	1.12	.83	1.58	6.53	36	.000**
<b>Syntactic complexity</b> MLT	Intervention	-.64	1.61	-1.15	-.13	-2.55	40	.015*
	Comparison	.02	.16	-.03	.07	.88	36	.386
<b>Lexical complexity</b> MSTTR_50	Intervention	-.00	.03	-.01	.01	-.66	40	.512
	Comparison	.00	.02	-.01	.01	.07	36	.947
<b>Fluency</b> WPT	Intervention	-8.24	29.88	-17.68	1.19	-1.77	40	.085
	Comparison	-1.49	24.59	-9.68	6.71	-.37	36	.715

Notes. EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text; \*  $p < .05$ ; \*\*  $p < .001$

ANCOVA results indicate no significant between-group difference in any of the four CAF indexes in Draft 2:  $F(1, 75) = 1.11, p = .296$  for accuracy,  $F(1, 75) = 1.91, p = 0.171$  for syntactic complexity;  $F(1, 75) = .49, p = .487$  in terms of lexical complexity, and  $F(1, 75) = 2.06, p = .155$  for fluency. In other words, when pre-existing differences in CAF indexes were controlled, the two groups performed similarly in each CAF index for Draft 2.

In summary, the treatment group and the comparison group made only a few significant changes in CAF indexes in Writing Task 6. Both the treatment group and the comparison

group made significant gains in accuracy between drafts; there was no significant difference between the two groups in accuracy for Draft 2. Only the treatment group made significant gains in syntactic complexity between drafts, there was no significant difference between the two groups in syntactic complexity for Draft 2. There was neither significant within-group difference, nor significant between-group difference, found for lexical complexity. Although the treatment group made gains in fluency between drafts and wrote more words in Draft 2 than the comparison group, none of the results were statistically significant.

### **4.3 Effects of peer feedback on writing performance**

This section reports comparisons of compositions written by the treatment group and the comparison group participants in the pre-test, post-test, and delayed post-test. After comparing baseline conditions of the two groups, mixed ANOVAs, repeated measure ANOVAs and independent t-tests were employed to analyze the three measures in scores and four measures in CAF indexes.

#### **4.3.1 Comparison of baseline conditions of the two groups**

The participants' compositions in the pre-test provided the baseline data for examining the effects of peer feedback on writing performance. As seen in Table 4.7 and Table 4.8, the two groups had similar writing proficiency prior to the intervention, as measured by the three writing scores and the four CAF indexes (*p* values all above .05).

Table 4.7 Descriptive statistics and one-way ANOVA analysis of scores in the pre-test

<b>Scores</b>	<b>Group</b>	<b>N</b>	<b>M</b>	<b>SD</b>	<b>F</b>	<b>df</b>	<b>p</b>
<b>Overall</b>	Treatment	41	68.85	5.20	.01	(1,76)	.924
	Comparison	37	68.97	5.87			
<b>Content</b>	Treatment	41	20.22	2.33	.01	(1,76)	.931
	Comparison	37	20.27	2.83			
<b>Organization</b>	Treatment	41	13.73	1.34	.75	(1,76)	.391
	Comparison	37	14.03	1.68			

Table 4.8 Descriptive statistics and one-way ANOVA analysis of CAF indexes in the pre-test

<b>CAF indexes</b>	<b>Group</b>	<b>N</b>	<b>M</b>	<b>SD</b>	<b>F</b>	<b>df</b>	<b>p</b>
<b>Accuracy</b> EP100W	Treatment	41	7.14	2.68	.59	(1,76)	.445
	Comparison	37	7.56	2.05			
<b>Syntactic complexity</b> MLT	Treatment	41	16.43	3.59	1.81	(1,76)	.183
	Comparison	37	15.42	2.98			
<b>Lexical complexity</b> MSTTR_50	Treatment	41	.76	.04	.73	(1,76)	.396
	Comparison	37	.76	.03			
<b>Fluency</b> WPT	Treatment	41	225.44	39.16	2.69	(1,76)	.105
	Comparison	37	242.16	50.72			

Notes. EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text

#### 4.3.2 Changes in writing scores

After the assumptions of normality, homogeneity, and sphericity were checked for the three writing scores, mixed ANOVAs were first applied to analyze their changes. One-

way repeated measure ANOVAs and independent t-tests were then used to investigate within-group differences and between-group differences.

#### 4.3.2.1 *Overall score*

Table 4.9 and Figure 4.1 display the descriptive results of overall scores in the pre-, post-, and delayed post-tests. As shown in Figure 4.1, overall scores for both groups increased in the post-test and delayed post-test, but the scores of the treatment group increased more than the scores of the comparison group in the post-test and the delayed post-test.

Table 4.9 Group means and standard deviations of overall scores in the pre-, post-, and delayed post-tests

<b>Group</b>	<b>N</b>	<b>Pre-test</b>		<b>Post-test</b>		<b>Delayed post-test</b>	
		<b>(T1)</b>		<b>(T2)</b>		<b>(T3)</b>	
		<b>M</b>	<b>SD</b>	<b>M</b>	<b>SD</b>	<b>M</b>	<b>SD</b>
Treatment	41	68.85	5.20	74.88	4.75	74.90	4.62
Comparison	37	68.97	5.87	71.95	4.92	71.43	5.10



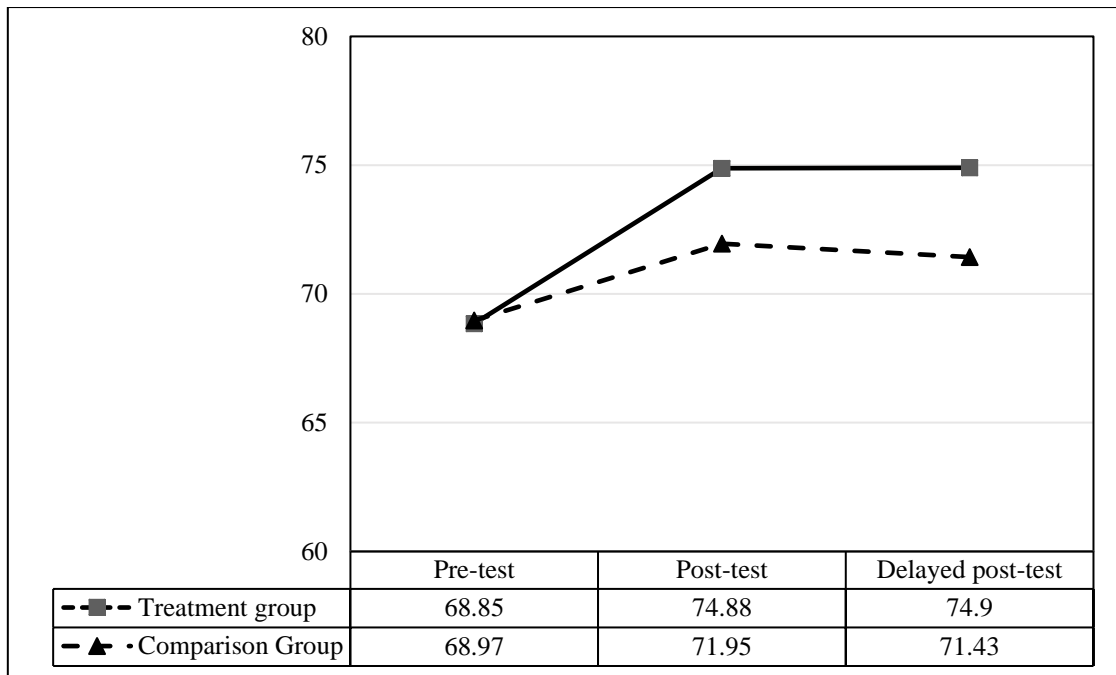


Figure 4.1 Mean overall scores in the pre-, post-, and delayed post-tests

The result of the mixed ANOVA for overall score is shown in Table 4.10. It reveals a significant main effect of time,  $F(2, 152) = 35.11, p < .001, \eta^2 = .316$ ; a significant main effect of group,  $F(1, 76) = 5.22, p = .025, \eta^2 = .064$ ; and a significant time  $\times$  group interaction effect,  $F(2, 152) = 35.11, p = .007, \eta^2 = .063$ .

Table 4.10 Effect of time and group condition on overall scores

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	35.11	(2,152)	.000**	.316
<b>Group</b>	5.22	(1,76)	.025*	.064
<b>Time <math>\times</math> group</b>	5.14	(2,152)	.007*	.063

Notes. \*  $p < .05$ ; \*\*  $p < .001$

One-way repeated measure ANOVA results indicate that the overall scores of both groups increased significantly over time (see Table 4.11). Participants in the treatment group made significant gains in overall scores in the post-test (Pre-M = 68.85, Post-M = 74.88,  $p < .001$ ,  $d = 1.13$ ), as well as in the delayed post-test (Pre-M = 68.85, Delayed-M = 74.90,  $p < .001$ ,  $d = 1.24$ ). Similarly, the comparison group made significant gains in overall scores in the post-test (Pre-M = 68.97, Post-M = 71.95,  $p = .021$ ,  $d = .47$ ), and in the delayed posttest (Pre-M = 68.97, Delayed-M = 71.43,  $p = .039$ ,  $d = .43$ ). While both groups' gains were retained in the delayed post-test, there were differences in the improvement of overall writing quality made over time. The treatment group had consistently large effect sizes, whereas the comparison group had medium effect sizes.

Table 4.11 Within-group comparisons in overall scores

Group	Pre vs. Post-		Post vs. Delayed		Pre vs. Delayed	
	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>
Treatment	.83	.000**	.75	1.000	.76	.000**
Comparison	1.04	.021*	.79	1.000	.94	.039*

Notes. Pre = Pre-test; Post = Post-test; Delayed = Delayed post-test; \*  $p < .05$ ; \*\*  $p < .001$

Table 4.12 shows that the treatment group had significantly better results than the comparison group in the post-test,  $p = .009$ , with a medium effect size,  $d = 0.61$ , as well as in the delayed post-test,  $p = .002$ , with a medium effect size,  $d = 0.71$ .

Table 4.12 Between-group comparisons in overall scores

	Group	N	M	SD	t	p	95% CI	
							LL	UL
Post-test	Treatment	41	74.88	4.75	2.68	.009*	.75	5.11
	Comparison	37	71.95	4.92				
Delayed post-test	Treatment	41	74.90	4.62	3.15	.002*	1.28	5.66
	Comparison	37	71.43	5.10				

Notes. \*  $p < .05$

To summarize, both the treatment group and the comparison group improved significantly in overall scores over time and retained their post-test gains in the delayed post-test. In addition, the treatment group significantly outperformed the comparison group in the post- and delayed post-tests.

#### 4.3.2.2 Content score

Table 4.13 and Figure 4.2 present the descriptive statistics for content scores of the two groups in the pre-, post-, and delayed post-tests. Whereas both groups' content scores increased in the post- and delayed post-tests, the treatment group outperformed the comparison group in the post- and delayed post-tests.

Table 4.13 Group means and standard deviations of content scores in the pre-, post-, and delayed post-tests

Group	N	Pre-test		Post-test		Delayed post-test	
		(T1)		(T2)		(T3)	
		M	SD	M	SD	M	SD
Treatment	41	20.22	2.33	22.05	2.43	22.15	2.33
Comparison	37	20.27	2.83	21.54	2.50	21.08	2.30

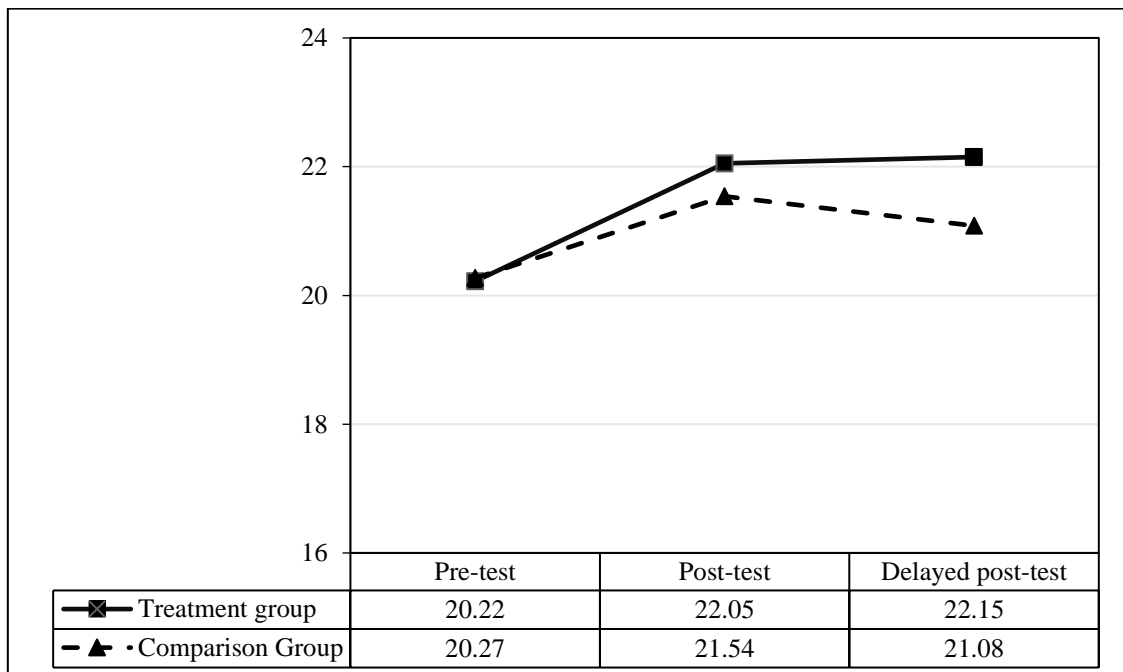


Figure 4.2 Mean content scores in the pre-, post-, and delayed post-tests

The result from the mixed ANOVA (see Table 4.14) shows that there is a significant main effect of time,  $F(2, 152) = 15.20, p < .001, \eta^2 = .167$ , but neither group condition nor the time  $\times$  group interaction effect is significant. As a result, only one-way repeated measure ANOVAs were conducted to investigate within-group differences.

Table 4.14 Effect of time and group condition on content scores

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	15.20	(2,152)	.000**	.167
<b>Group</b>	1.40	(1,76)	.240	.018
<b>Time × group</b>	1.65	(2,152)	.196	.021

Notes. \*\*  $p < .001$

One-way repeated measure ANOVA results, as displayed in Table 4.15, reveal that only the treatment group increased significantly in content scores over time. They made significant gains in content scores in the post-test (Pre-M=20.22; Post-M =22.05;  $p < .001$ ,  $d = .71$ ) and in the delayed post-test (Pre-M=20.22; Delayed-M=22.15;  $p < .001$ ,  $d = .76$ ). However, the content scores of the comparison group did not vary significantly over the three time points (Pre- vs. Post-tests,  $p = .076$ ; Post- vs. Delayed post-tests,  $p = .735$ ).

Table 4.15 Within-group comparisons in content scores

<b>Group</b>	<b>Pre- vs. Post-</b>		<b>Post- vs. Delayed</b>		<b>Pre vs. Delayed</b>	
	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>
Treatment	.40	.000**	.38	1.000	.40	.000**
Comparison	.55	.076	.39	.735	.49	.326

Notes. Pre = Pre-test; Post = Post-test; Delayed = Delayed post-test; \*\*  $p < .001$

To sum up, only the treatment group increased significantly in content scores over time and retained the gains in the delayed post-test. As regards the between-group differences,

although the treatment group had higher content scores than the comparison group in the post- and delayed post-tests, neither difference was statistically significant.

#### 4.3.2.3 *Organization score*

Table 4.16 and Figure 4.3 present the descriptive statistics of organization scores of the two groups in the pre-, post-, and delayed post-tests. As can be seen from Figure 4.3, both groups increased in organization scores in the post-test and delayed post-test; the treatment group made greater gains than the comparison group in the post-test and delayed post-test.

Table 4.16 Group means and standard deviations of organization scores in the pre-, post-, and delayed post-tests

Group	N	Pre-test (T1)		Post-test (T2)		Delayed post-test (T3)	
		M	SD	M	SD	M	SD
Treatment	41	13.73	1.34	15.59	1.41	15.22	1.31
Comparison	37	14.03	1.68	14.62	1.52	14.54	1.64

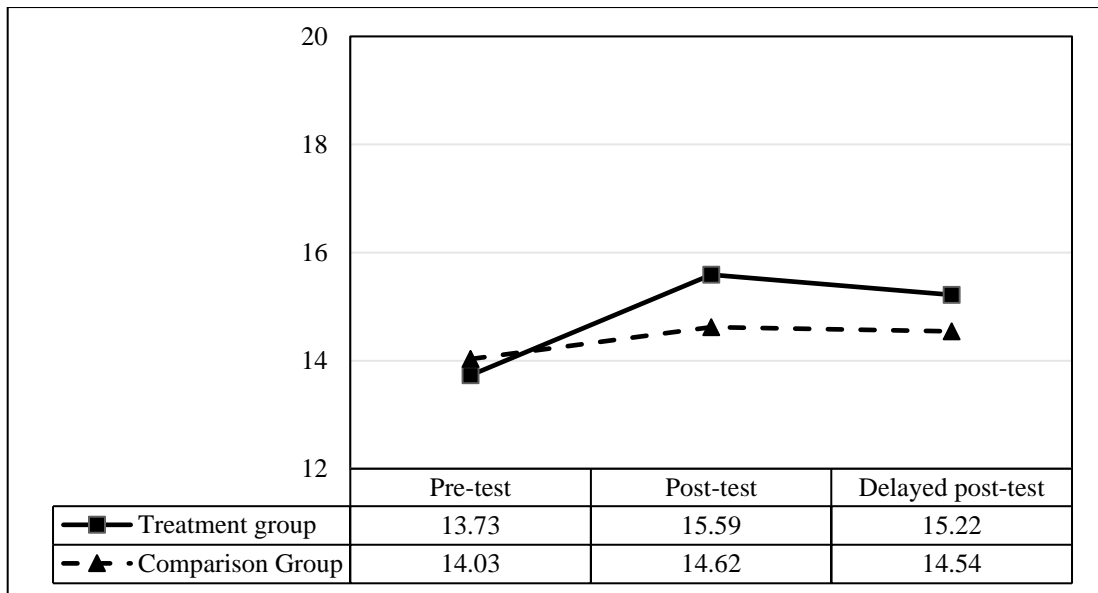


Figure 4.3 Mean organization scores in the pre-, post-, and delayed post-tests

Table 4.17 presents the mixed ANOVA result, indicating a significant main effect of time,  $F(2, 152) = 25.82, p < .001, \eta^2 = .254$ , and a significant time  $\times$  group interaction effect,  $F(2, 152) = 6.62, p = .002, \eta^2 = .080$ . However, no significant main effect of group was observed,  $F(1, 76) = 2.87, p = .094, \eta^2 = .036$ .

Table 4.17 Effect of time and group condition on organization scores

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	25.82	(2,152)	.000**	.254
<b>Group</b>	2.87	(1,76)	.094	.036
<b>Time <math>\times</math> group</b>	6.62	(2,152)	.002*	.080

Notes. \*  $p < .05$ ; \*\*  $p < .001$

One-way repeated measure ANOVA results indicate that only the treatment group increased significantly in organization scores over time (see Table 4.18). They made significant gains, with consistently large effect sizes, in organization scores in the post-test (Pre-M = 13.73, Post-M =15.59,  $p < .001$ ,  $d = 1.25$ ), as well as in the delayed post-test (Pre-M =13.73, Delayed-M =15.22,  $p < .001$ ,  $d = 1.10$ ). The comparison group did not vary in organization scores significantly over the three time points (Pre- vs. Post-tests,  $p = .222$ ; Post- vs. Delayed post-tests,  $p = .244$ ).

Table 4.18 Within-group comparisons in organization scores

Group	Pre vs. Post		Post vs. Delayed		Pre vs. Delayed	
	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>
Treatment	.23	.000**	.24	1.000	.21	.000**
Comparison	.32	.222	.25	1.000	.29	.244

*Notes.* Pre = Pre-test; Post = Post-test; Delayed = Delayed post-test; \*\*  $p < .001$

Independent t-tests were conducted to investigate the between-group differences in organization scores. As shown in Table 4.19, the treatment group had significantly higher organization scores than the comparison group in the post-test,  $p = .005$ , with a medium effect size,  $d = 0.65$ , as well as in the delayed post-test,  $p = 0.046$ , with a small effect size,  $d = 0.46$ .



Table 4.19 Between-group comparisons in organization scores

	Group	N	M	SD	t	p	95% CI	
							LL	UL
Post-test	Treatment	41	15.59	1.41	2.87	.005*	.30	1.63
	Comparison	37	14.62	1.55				
Delayed post-test	Treatment	41	15.22	1.31	2.03	.046*	.01	1.35
	Comparison	37	14.54	1.64				

Notes. \*  $p < .05$ ; \*\*  $p < .001$

In summary, only the treatment group increased significantly in organization scores over time with gains retained in the delayed post-test. The treatment group also exhibited significantly higher scores than the comparison group in both the post-test and the delayed post-test.

### 4.3.3 Changes in CAF indexes

As assumptions of normality, homogeneity, and sphericity were met for the three CAF indexes (EP100W, MLT and MSTTR\_50), mixed ANOVAs were applied, followed by one-way repeated measure ANOVAs and independent t-tests. For the fluency index of WPT, because the assumption of normality was met but the assumptions of homogeneity and sphericity were violated, one-way repeated measure ANOVAs and independent t-tests were directly applied.

#### 4.3.3.1 Accuracy

As displayed in Table 4.20 and Figure 4.4, whereas the number of errors (per 100 words) decreased in the post- and the delayed post-tests for the treatment group, the number of

errors (per 100 words) for the comparison group decreased in the post-test, but then increased to almost its pre-test level in the delayed post-test.

Table 4.20 Group means and standard deviations of EP100W in the pre-, post-, and delayed post-tests

Group	N	Pre-test (T1)		Post-test (T2)		Delayed post-test (T3)	
		M	SD	M	SD	M	SD
		Treatment	41	7.14	2.68	5.54	2.34
Comparison	37	7.56	2.05	6.42	2.23	7.54	2.29

Notes. EP100W = Errors per 100 words

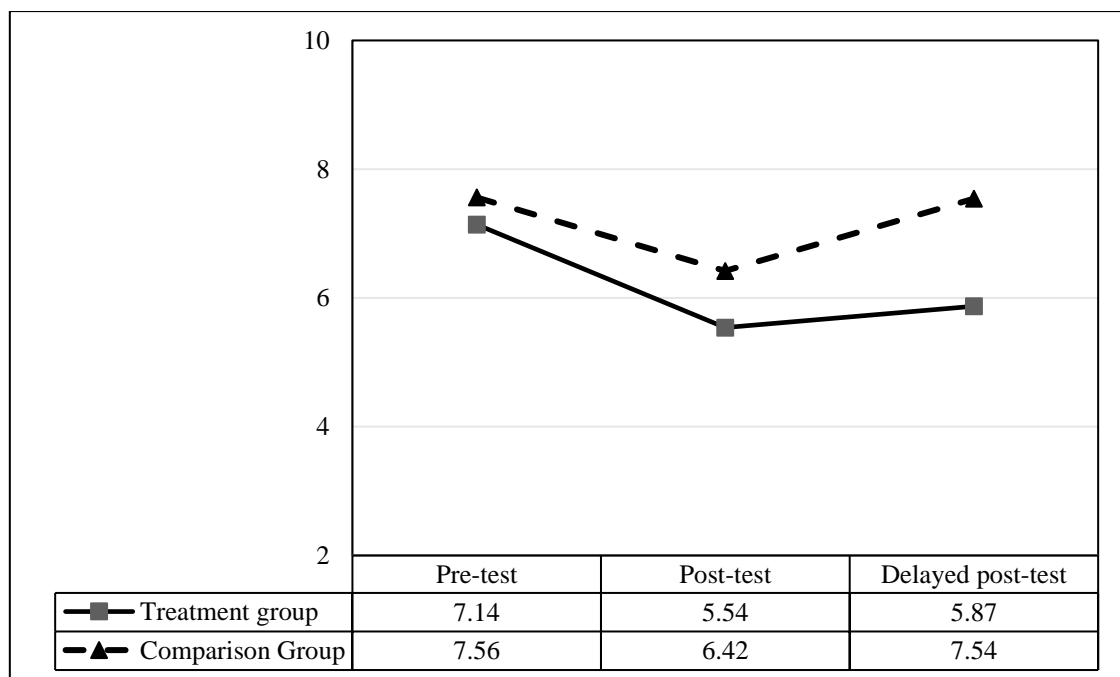


Figure 4.4 The mean number of EP100W in the pre-, post-, and delayed post-tests

Table 4.21 presents the mixed ANOVA result for the accuracy index, EP100W. There was a significant main effect of time,  $F(2, 152) = 11.84, p < .001, \eta^2 = .135$ , and a significant main effect of group,  $F(1, 76) = 5.66, p = .020, \eta^2 = .069$ . However, the time  $\times$  group interaction effect was not significant,  $F(2, 152) = 2.48, p = .087, \eta^2 = .032$ .

Table 4.21 Effect of time and group condition on EP100W

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	11.84	(2,152)	.000**	.135
<b>Group</b>	5.66	(1,76)	.020*	.069
<b>Time <math>\times</math> group</b>	2.48	(2,152)	.087	.032

*Note.* EP100W = Errors per 100 words; \*  $p < .05$ ; \*\*  $p < .001$

One-way repeated measure ANOVA results indicate that both groups changed significantly in the number of EP100W over time (see Table 4.22). The number of EP100W for the treatment group decreased significantly in the post-test (Pre-M = 7.14, Post-M = 5.54,  $p = .001, d = .61$ ), as well as in the delayed post-test (Pre-M = 7.14, Delayed-M = 5.87,  $p = .011, d = .48$ ). For the comparison group, the number of EP100W decreased significantly in the post-test (Pre-M = 7.56, Post-M = 6.42,  $p = .014, d = .49$ ), but increased significantly in the delayed post-test (Post-M = 6.42, Delayed-M = 7.54,  $p = 0.013, d = .50$ ), restoring its level to that of the pre-test (Pre-M = 7.56, Delayed-M = 7.54,  $p = 1.000$ ).

Table 4.22 Within-group comparisons in EP100W

	Pre vs. Post		Post vs. Delayed		Pre vs. Delayed	
	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>
Treatment	.41	.001*	.41	1.000	.41	.011*
Comparison	.38	.014*	.37	.013*	.40	1.000

*Notes.* EP100W = Errors per 100 words; Pre = Pre-test; Post = Post-test; Delayed = Delayed post-test; \*  $p < .05$

Table 4.23, the results of between-group comparisons in the number of EP100W in the post- and delayed post-tests, show that the treatment group significantly outperformed the comparison group in accuracy only in the delayed post-test. In the post-test, there was no significant difference in the number of EP100W between the two groups,  $p = .095$ ; in the delayed post-test, however, the difference in the number of EP100W between the two groups was significant,  $p = .002$ , with a medium effect size,  $d = 0.73$ .

Table 4.23 Between-group comparisons in EP100W

	Group	N	M	SD	t	<i>p</i>	95% CI	
							LL	UL
Post-test	Treatment	41	5.54	2.34	-1.69	.095	-1.91	.16
	Comparison	37	6.42	2.23				
Delayed post-test	Treatment	41	5.87	2.29	-3.21	.002*	-2.70	-.63
	Comparison	37	7.54	2.28				

*Notes.* EP100W = Errors per 100 words; \*  $p < .05$

To summarize, both the treatment group and the comparison group improved significantly in accuracy in the post-test, but only the treatment group retained the improvement in the delayed post-test. The treatment group significantly outperformed the comparison group in accuracy only in the delayed post-test.

#### 4.3.3.2 *Syntactic complexity*

Table 4.24 and Figure 4.5 present the descriptive statistics for the syntactic complexity index, MLT, in the pre-, post-, and delayed post-tests for the treatment group and the comparison group. As visually depicted in Figure 4.5, the treatment group decreased in MLT over time, whereas the comparison group increased in MLT over time.

Table 4.24 Group means and standard deviations of MLT in the pre-, post-, and delayed post-tests

Group	N	Pre-test (T1)		Post-test (T2)		Delayed post-test (T3)	
		M	SD	M	SD	M	SD
Treatment	41	16.43	3.59	16.02	2.73	16.01	2.71
Comparison	37	15.42	2.98	15.85	2.78	15.87	3.65

*Notes. MLT = Mean length of T-unit*

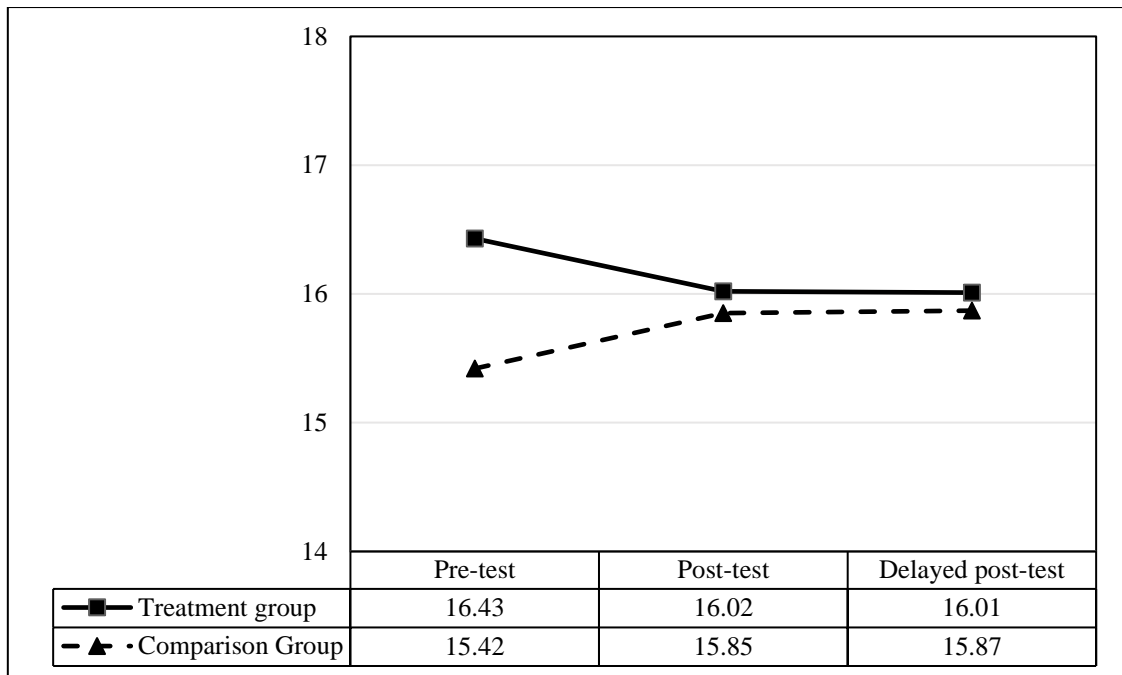


Figure 4.5 The MLT in the pre-, post-, and delayed post-tests

The Mixed ANOVA result shows that there was neither a significant effect of time,  $F(2, 152) = .00, p = .999, \eta^2 = .000$ , nor a significant main effect of group,  $F(1, 76) = .78, p = .379, \eta^2 = .010$ . The time  $\times$  group interaction effect was also not significant,  $F(2, 152) = .66, p = .515, \eta^2 = .009$  (see Table 4.25). Neither within-group nor between-group difference, was therefore examined.

Table 4.25 Effect of time and group condition on MLT

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	.00	(2,152)	.999	.000
<b>Group</b>	.78	(1,76)	.379	.010
<b>Time × group</b>	.66	(2,152)	.515	.009

*Notes.* MLT = Mean length of T-unit

In brief, although the treatment group decreased in MLT and the comparison group increased in MLT over time, neither within-group differences nor between-group differences were statistically significant.

#### 4.3.3.3 *Lexical complexity*

Table 4.26 and Figure 4.6 present the descriptive statistics for the lexical index, MSTTR<sub>50</sub>, in the pre-, post-, and delayed post-tests for the treatment group and the comparison group. As seen in the table, lexical complexity for the treatment group and the comparison group changed little over time.

Table 4.26 Group means and standard deviations of MSTTR\_50 in the pre-, post-, and delayed post-tests

Group	N	Pre-test		Post-test		Delayed post-test	
		(T1)		(T2)		(T3)	
		M	SD	M	SD	M	SD
Treatment	41	.76	.04	.76	.04	.75	.03
Comparison	37	.76	.03	.77	.03	.75	.04

Notes. MSTTR\_50 = Mean segmental type token ratio per 50 words

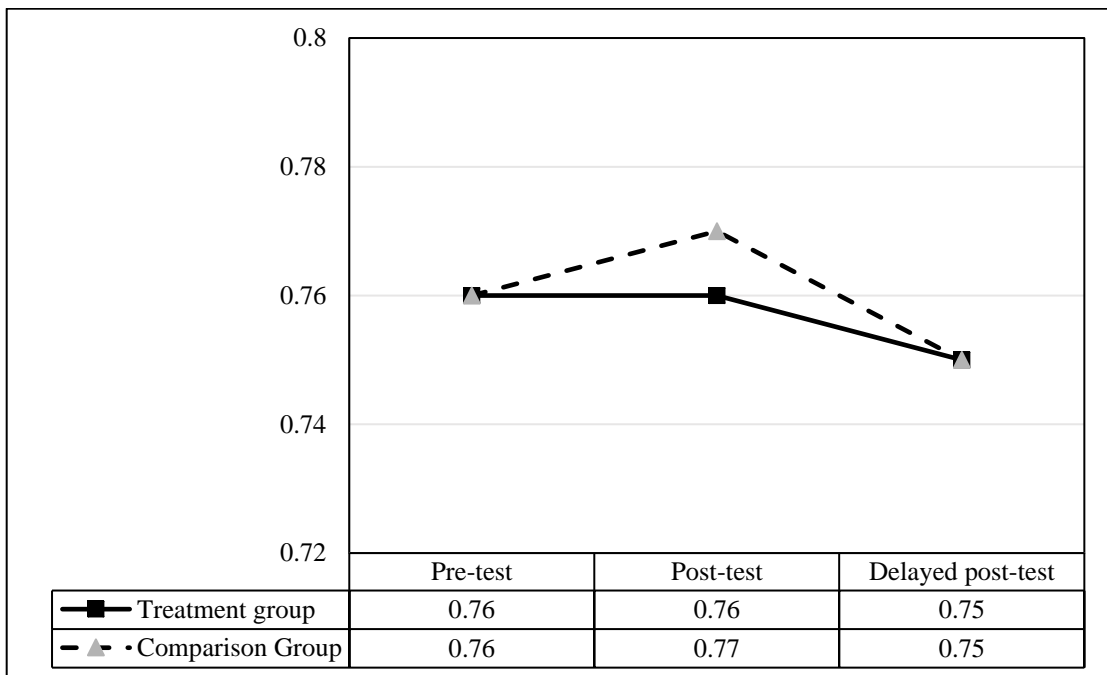


Figure 4.6 The MSTTR\_50 in the pre-, post-, and delayed post-tests

The mixed ANOVA result shows no significant main effect of time,  $F(2, 152) = 2.81$ ,  $p = .064$ ,  $\eta^2 = .036$ ; no significant main effect of group,  $F(1, 76) = .91$ ,  $p = .342$ ,  $\eta^2 = .012$ ; and no significant time  $\times$  group interaction effect,  $F(2, 152) = .05$ ,  $p = .951$ ,



$\eta^2 = .001$  (see Table 4.27). Neither within-group nor between-group difference, was therefore examined.

Table 4.27 Effect of time and group condition on MSTTR\_50

	<i>F</i>	<i>df</i>	<i>p</i>	partial $\eta^2$
<b>Time</b>	2.81	(2,152)	.064	.036
<b>Group</b>	.91	(1,76)	.342	.012
<b>Time <math>\times</math> group</b>	.05	(2,152)	.951	.001

*Notes.* MSTTR\_50 = Mean segmental type token ratio per 50 words

To sum up, the treatment group and the comparison group had similar performance in lexical complexity over time, with neither within-group differences nor between-group differences

#### 4.3.3.4 *Fluency*

Table 4.28 and Figure 4.7 present the descriptive statistics for the fluency index, WPT, for the two groups in the pre-, post-, and delayed post-tests. As shown in Figure 4.7, the text length of the treatment group increased in the post-test, but then decreased to its pre-test level in the delayed post-test; the text length of the comparison group decreased over time.

Table 4.28 Group means and standard deviations for WPT in the pre-, post-, and delayed post-tests

Group	N	Pre-test (T1)		Post-test (T2)		Delayed post-test (T3)	
		M	SD	M	SD	M	SD
		Treatment	41	225.44	39.16	252.22	42.98
Comparison	37	242.16	50.72	233.27	41.55	226.35	51.79

Notes. WPT = Words per text

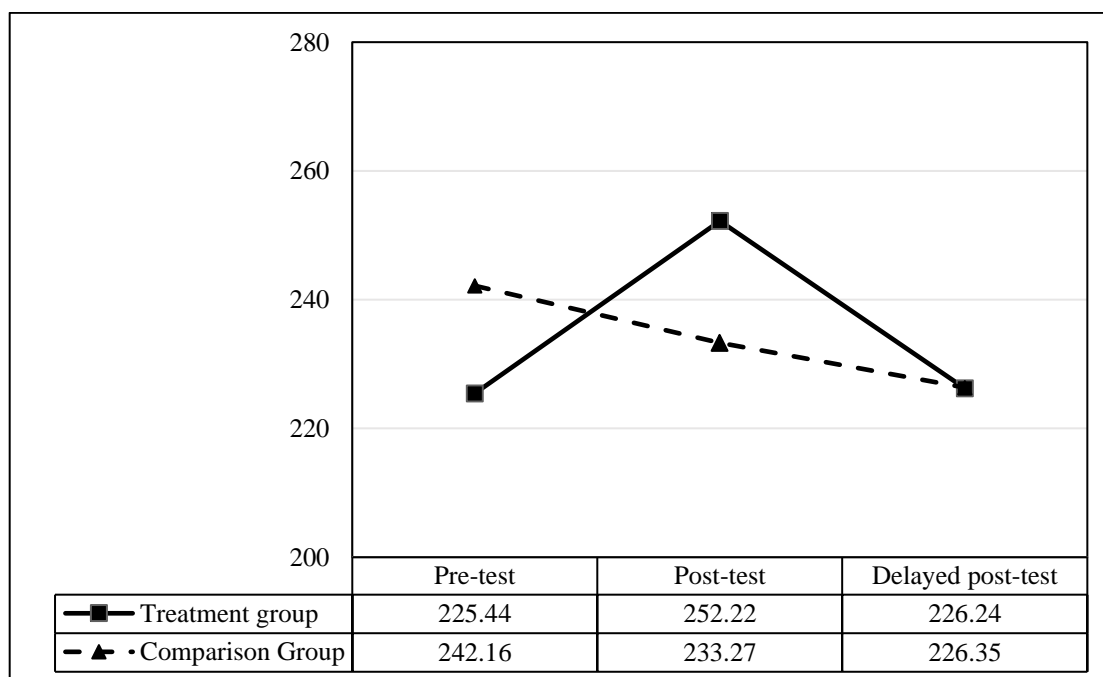


Figure 4.7 The WPT in the pre-, post-, and delayed post-tests

As mentioned in 4.3.3, one-way repeated measure ANOVAs and independent t-tests were used to investigate within-group differences and between-group differences in WPT, the fluency index, respectively. One-way repeated measure ANOVA results, as displayed in Table 4.29, indicate that the text length of the treatment group increased

significantly in the post-test (Pre-M= 225.44; Post-M = 252.22;  $p = .009$ ;  $d = .49$ ), and then decreased significantly in the delayed post-test (Post-M = 252.22; Delayed-M = 226.24;  $p < .001$ ;  $d = .76$ ), returning to its level in the pre-test (Pre-M<sub>INT</sub> = 225.44; Delayed-M<sub>INT</sub> = 226.24,  $p = 1.000$ ). The text length of the comparison group, however, did not change significantly over time (Pre- vs. Post-tests,  $p = .909$ ; Post- vs. Delayed post-tests,  $p = 1.000$ ).

Table 4.29 Within-group comparisons in WPT

Group	Pre vs. Post		Post vs. Delayed		Pre vs. Delayed	
	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>	<i>SE</i>	<i>p</i>
Treatment	8.50	.009*	5.38	.000**	7.75	1.000
Comparison	8.51	.909	8.08	1.000	6.82	.079

*Notes.* WPT = Words per text; \*  $p < .05$ ; \*\*  $p < .001$

Table 4.30 shows the results of between-group comparisons in WPT in the post- and delayed post-tests. There was no significant difference in WPT between the two groups either in the post-test ( $p = .052$ ) or in the delayed post-test ( $p = .992$ ). It should be noted that in the post-test, the WPT for the treatment group were greater than that for the comparison group (252.22 vs. 233.27).

Table 4.30 Between-group comparisons in WPT

	Group	N	M	SD	t	p	95% CI	
							LL	UL
Post-test	Treatment	41	252.22	42.98	1.98	.052	-.16	38.06
	Comparison	37	233.27	41.55				
Delayed post-test	Treatment	41	226.24	37.30	-.01	.992	-20.32	20.10
	Comparison	37	226.35	51.80				

*Notes.* WPT = Words per text

To summarize, the text length of only the treatment group participants increased significantly in the post-test but returned to their pre-test level in the delayed post-test. Although the treatment group produced longer texts than the comparison group in the post-test, there was no significant difference in text length between the two groups either in the post- or delayed post-tests.

#### 4.4 Chapter summary

This chapter has reported the effects of peer feedback on text revisions and writing performance. The effects of peer feedback on text revisions are summarized in Table 4.31. In overall, content and organization scores, both the treatment group and the comparison group made significant gains between drafts, but the treatment group significantly outperformed the comparison group in their second drafts. The treatment group and the comparison group made only a few significant changes in language use. Although the accuracy of both groups improved significantly between drafts, the two groups performed similarly in their second drafts. Only the treatment group made significant gains in syntactic complexity between drafts but the two groups performed

similarly in their second drafts. Neither significant within-group differences nor significant between-group differences were found in lexical complexity and fluency.

Table 4.31 Changes in writing scores and CAF indexes in Writing Task 6

Measure	Within-group		Between-group
	Treatment group	Comparison group	
<b>Overall score</b>	D1 < D2 <i>p</i> = .000**	D1 < D2, <i>p</i> = .000**	TRE > COM <i>p</i> = .015*
<b>Content score</b>	D1 < D2 <i>p</i> = .000**	D1 < D2 <i>p</i> = .045*	TRE > COM <i>p</i> = .020*
<b>Organization score</b>	D1 < D2 <i>p</i> = .000**	D1 < D2 <i>p</i> = .000**	TRE > COM <i>p</i> = .001*
<b>Language</b>			
Accuracy: EP100W	D1 > D2 <i>p</i> = .000**	D1 > D2 <i>p</i> = .000**	TRE < COM <i>p</i> = .296
Syntactic complexity: MLT	D1 < D2 <i>p</i> = .015*	D1 < D2 <i>p</i> = .386	TRE < COM <i>p</i> = .171
Lexical complexity: MSTTR_50	D1 = D2 <i>p</i> = .512	D1 < D2 <i>p</i> = .947	TRE = COM <i>p</i> = .487
Fluency: WPT	D1 < D2 <i>p</i> = .085	D1 < D2 <i>p</i> = .715	TRE > COM <i>p</i> = .155

*Notes.* EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text; TRE = Treatment group; COM = Comparison group; \**p* < .05; \*\**p* < .001

The effects of peer feedback on writing performance are summarized in Table 4.32. In terms of the overall quality, both the treatment group and the comparison group improved significantly in the post-test and retained the effect in the delayed post-test;

the treatment group significantly outperformed the comparison group in the post-test and delayed post-test. Only the treatment group made significant gains in content scores and organization scores over time and retained the gains in the delayed post-test. Whereas there was no significant difference in content scores between the two groups in the post- or delayed post-tests, there was significant differences in organization scores between the two groups in the post- and delayed post-tests. Both the treatment group and the comparison group improved significantly in accuracy in the post-test, but only the treatment group retained the improvement in the delayed post-test. The treatment group also significantly outperformed the comparison group in accuracy, but only in the delayed post-test. For syntactic complexity and lexical complexity, neither within-group nor between-groups difference was statistically significant. With fluency, only the treatment group participants increased significantly in text length in the post-test but returned to their pre-test level in the delayed post-test; there was no significant difference in text length between the two groups either in the post- or delayed post-tests.

Table 4.32 Changes in writing scores and CAF indexes over time

		Within-group differences			Between-group differences	
		T1 vs. T2	T2 vs. T3	T1 vs. T3	T2	T3
<b>Overall quality</b>	TRE	T1 < T2, $p < .001^{**}$	T2 < T3, $p = 1.000$	T1 < T3, $p < .001^{**}$	TRE > COM	TRE > COM
	COM	T1 < T2, $p = .021^*$	T2 > T3, $p = 1.000$	T1 < T3, $p = .039^*$	$p = .009^*$	$p = .002^*$
<b>Content</b>	TRE	T1 < T2, $p < .001^{**}$	T2 < T3, $p = 1.000$	T1 < T3, $p < .001^{**}$		
	COM	T1 < T2, $p = .076$	T2 > T3, $p = .735$	T1 < T3, $p = .326$		
<b>Organization</b>	TRE	T1 < T2, $p < .001^{**}$	T2 > T3, $p = 1.000$	T1 < T3, $p < .001^{**}$	TRE > COM	TRE > COM
	COM	T1 < T2, $p = .222$	T2 > T3, $p = 1.000$	T1 < T3, $p = .244$	$p = .005^*$	$p = .046^*$
<b>Language</b>						
Accuracy: EP100W	TRE	T1 > T2, $p = .001^*$	T2 < T3, $p = 1.000$	T1 > T3, $p = .011^*$	TRE < COM	TRE < COM
	COM	T1 > T2, $p = .014^*$	T2 < T3, $p = .013^*$	T1 > T3, $p = 1.000$	$p = .095$	$p = .002^*$
Syntactic complexity: MLT	TRE					
	COM					
Lexical complexity: MSTTR_50	TRE					
	COM					
Fluency: WPT	TRE	T1 < T2, $p = .009^*$	T2 > T3, $p < .001^{**}$	T1 < T3, $p = 1.000$	TRE > COM	TRE > COM
	COM	T1 > T2, $p = .909$	T2 > T3, $p = 1.000$	T1 > T3, $p = .079$	$p = .052$	$p = .992$

Notes. EP100W = Errors per 100 words; MLT = Mean length of T-unit, MSTTR\_50 = Mean segmental type token ratio per 50 words; WPT = Words per text; T1 = Pre-test; T2 = Post-test; T3 = Delayed post-test; TRE = Treatment group; COM = Comparison group;  $*p < .05$ ;  $**p < .001$

## Chapter 5 Case Study Findings

### 5.1 Chapter overview

This chapter presents the findings for the third research question which investigates the processing of peer feedback in revision.

RQ 3 How do Chinese EFL learners process peer feedback when writing their second drafts?

*What cognitive operations do Chinese EFL learners employ when processing peer feedback?*

*What approaches do Chinese EFL learners take when processing peer feedback?*

A multiple case study was employed to examine how the six case study participants used received feedback points when writing their second drafts. Data collected from Writing Task 6 included think-aloud protocols, transcripts of paired discussions on feedback as well as their two drafts and the accompanying PF forms.

There are four sections in this chapter. Following the chapter overview, the findings about feedback points that the case study participants received are presented. The following section reports how the case study participants processed peer feedback when writing their second drafts. The final section summarizes findings in this chapter.

Before the case study findings are presented, the six case study participants are briefly described. They were selected from the treatment group based on two criteria: demonstrating the ability in performing think-aloud protocols and having sufficient time



to attend the think-aloud training and the two think-aloud sessions. Demographic information about the six case study participants is presented in Table 5.1. A pseudonym was assigned to each participant to ensure anonymity.

Table 5.1 Case study participants' demographic information

<b>Case</b>	<b>Gender</b>	<b>Age</b>	<b>Major</b>
Yang	Male	19	Electrical Science
Sun	Female	18	Chemistry
Shen	Male	19	Electrical Science
Zhang	Female	19	Chemistry
Liu	Female	19	Biology
Wu	Male	18	Physics

## **5.2 Received peer feedback points**

The six case study participants received a total of 100 feedback points on their first drafts of Writing Task 6. Table 5.2 shows the categorization of feedback points in terms of focus and validity, and that, although the number of feedback points that each case study participant received varied, their quality was generally satisfactory. While each case study participant received feedback on content, organization, and language use, there were more feedback points on language use than on content and organization, with only a small percentage (12%) of feedback points judged to be invalid.

Table 5.2 Peer feedback points received on Draft 1 of Writing Task 6

	<b>Yang</b>	<b>Sun</b>	<b>Shen</b>	<b>Zhang</b>	<b>Liu</b>	<b>Wu</b>	<b>Total</b>
<b>Feedback points</b>	7	21	21	14	24	13	100
<b>Feedback focus</b>							
Content	1	3	2	6	6	2	20 (20%)
Organization	1	5	4	1	3	4	18 (18%)
Language	5	13	15	7	15	7	62 (62%)
<b>Feedback Validity</b>							
Valid	6	19	19	13	22	9	88 (88%)
Invalid	1	2	2	1	2	4	12 (12%)

### 5.3 Processing of peer feedback in revision

The findings reported in this section are based upon the 82 PF-triggered revision episodes, identified from the think-aloud protocols, in which the case study participants attended to peer feedback points when writing their second drafts (see Table 5.3). In the following two sub-sections, the nine cognitive operations and two approaches used by the case study participants when processing peer feedback points are presented and illustrated with representative excerpts from the think-aloud protocols.

Table 5.3 Peer feedback points received and attended to in revision

	<b>Yang</b>	<b>Sun</b>	<b>Shen</b>	<b>Zhang</b>	<b>Liu</b>	<b>Wu</b>	<b>Total</b>
<b>Feedback points received</b>	7	21	21	14	24	13	100
<b>Feedback points attended to in revision</b>	7	15	16	12	22	10	82

### 5.3.1 Cognitive operations

The 82 PF-triggered revision episodes included a total of 412 cognitive operations in nine categories (see Table 5.4). *Referring to PF* and *generating changes/no change* were most frequently employed, as they were present in each PF-triggered revision episode. The cognitive operations of *resourcing*, *retrieving L2 knowledge*, *monitoring* and *justifying* were also frequently employed. In the following sub-sections, each cognitive operation will be illustrated with excerpts from students' think-aloud protocols.

Table 5.4 The cognitive operations and their frequencies in the 94 PF-triggered revision episodes

Cognitive operation	Frequency
Referring to PF	82
Referring to text	15
Retrieving L2 knowledge	36
Justifying	31
Resourcing	61
Translating	19
Monitoring	33
Evaluating	28
Generating changes/no change	107
Total cognitive operations	412

#### 5.3.1.1 *Referring to PF*

A cognitive operation, in which the case study participants referred to a feedback point that they had received, was the starting point of each PF-triggered revision episode. In most cases, they read aloud or paraphrased a written feedback point in the context of its relevant text segment, as can be seen from the following excerpt.

#### **Excerpt 5.1**

(Sun, PF-triggered revision episode 4)

[Reading a sentence in her first draft] “**Besides, we are also free to learn how the literature attracts its readers and expresses its thesis with its unique**

**way of organizing its language.”** My peer marked **why** at the end of this sentence. He told me this sentence was too abrupt.

In other instances, the case study participants read aloud or paraphrased a written feedback point or referred to what was explained or negotiated about it in discussions, as shown in the following two excerpts.

### **Excerpt 5.2**

(Liu, PF-triggered revision episode 1)

My peer told me that I should not use a colon after “**saying**” [in the sentence “As we can see, there is a mother sending a book to her son and saying: ‘Just think of it as you are reading a long text-message’ in the picture”].

### **Excerpt 5.3**

(Zhang, PF-triggered revision episode 1)

My peer told me that my description was too detailed. She told me not to use the exact words in the directions.

(Discussion between Zhang and her peer)

**Peer:** Your description is too detailed. I remembered that for IELTS writing, you cannot use the words in the directions without making any change. You can summarize your ideas in a shorter sentence.

**Zhang:** OK. Thanks.

Similarly, when Shen came to the sentence in his first draft “That requires a few and will pay more than you think back”, which his peer did not comment on through written feedback but they jointly corrected in the discussion, he mentioned the changes that they agreed upon, “My peer told me to revise the sentence into ‘**That requires much but will pay back more**’” (Shen, PF-triggered revision episode 13).

#### 5.3.1.2 *Referring to text*

Similar to *referring to PF*, *referring to text* is a cognitive operation in which the case study participants read aloud or paraphrased some part(s) of their drafts. The following excerpt was from Liu, in her response to the feedback of adding a phrase or sentence to link Paragraphs 3 and Paragraph 4.

#### **Excerpt 5.4**

(Liu, PF-triggered revision episode 16)

Paragraph 3 was about the importance of reading. If you don’t read, you will make no progress. In Paragraph 4, I wrote we should read rather than do other meaningless things.

Liu referred to Paragraph 3 and Paragraph 4 in her first draft by summarizing main ideas for each paragraph. In this example, *referring to text* stimulated Liu to generate ideas for the added sentence, “Now that we have known the importance of reading, appropriate action can be taken” in her second drafts.

It is noteworthy that the participants occasionally employed the cognitive operation of *referring to text* when dealing with feedback on content and organization. In some

instances, rather than reading aloud or paraphrasing several sentences, they directly read aloud the thesis statement or topic sentences in body paragraphs.

### 5.3.1.3 *Retrieving L2 knowledge*

As the writing task was in English, the case study participants naturally relied on their L2 knowledge when revising. The think-aloud data indicate that they retrieved L2 words and grammatical structures when carrying out revisions triggered by peer feedback.

Of the six case study participants, Sun retrieved L2 words and expressions most frequently. The cognitive operation of *retrieving L2 knowledge* was present in 10 of her 15 PF-triggered revision episodes. By relying on her vocabulary knowledge, she identified L2 words and idioms as well as meanings of specific L2 words. She retrieved 11 words and expressions in English (“horizon”, “so as to”, “adopt”, “precisely”, “at the same time”, “in addition”, “rapid scan”, “detailed”, “discover”, “out of sight”, and “on a regular basis”), and used them in her second draft. The following excerpt from Sun illustrates how she retrieved a L2 word to use in her second draft.

#### **Excerpt 5.5**

(Sun, PF-triggered revision episode 1)

“**Horizon**”, 人生视野 [meaning “horizon”], is a noun. I should use “**broaden one’s horizons**”, 拓宽视野 [meaning “to broaden one’s horizons”].

The retrieval of L2 words occurred when Sun responded to the feedback asking her to modify “enriched values” in her thesis statement, “In pursuit of better writing ability and enriched values, we’re supposed to read literature devotedly and frequently.” Because her peer gave only the Chinese equivalent “人生视野” [meaning “horizon”],

Sun retrieved the related information of an L2 word from her long-term memory, including its spelling and part of speech. Notably, she also recalled “broaden our horizons” as a collocation. In her second draft, she wrote “Yet in order to develop writing ability and broaden our horizons, people are supposed to read literature devotedly and frequently.”

Moreover, case study participants consciously drew on their grammatical knowledge when implementing revisions. For instance, when Yang was working on shortening his description of a picture, he used his prior knowledge about nominalizations, “I can use ‘**the action of**’ [to revise segment ‘here’s a teacher handing a book to a little boy’ into ‘the teacher’s action of handing a book to the little boy’]” (Yang, PF-triggered revision episode 1). Similarly, Zhang activated her knowledge of subordinate clauses, “Can I use one sentence to summarize the information in the picture? Probably I can include some clauses” (Zhang, PF-triggered revision episode 1). Other instances of L2 knowledge retrieval can be found in excerpts exemplifying the cognitive operation of *justifying*, in which the participants retrieved L2 grammatical knowledge and L2 writing knowledge to justify their choice (see 5.3.1.4).

There are, however, a few instances, in which the case study participants made inappropriate revisions based on their incorrect or incomplete L2 knowledge. For instance, Zhang used a misspelled word, “outsides”, in her second draft. The “completely” in Shen’s revised sentence “In fact, reading literature provides knowledge completely and teaches a person to be persistent till something is finished” was also inappropriately used, even though he had consulted online dictionaries before incorporating it in his second draft.



#### 5.3.1.4 *Justifying*

*Justifying*, the cognitive operation of providing reasons, was evident in PF-triggered revision episodes, as the case study participants justified the feedback they received as well as the changes made in their second drafts.

The following excerpt illustrates how Shen justified a feedback point before implementing corresponding changes.

##### **Excerpt 5.6**

(Shen, PF-triggered revision episode 4)

*I used “**on the one hand**” to introduce the idea of developing thinking ability, and “**on other one hand**” to introduce the idea of developing one’s perseverance. These two ideas were not opposite, so the two linking phrases were inappropriate.*

This excerpt was taken from a revision episode in which Shen processed his peer’s feedback that “‘**on the one hand**’ and ‘**on the other hand**’ are used to introduce two opposite points” (Shen’s PF form). He then explained why the two transition phrases were inappropriately used.

In addition, the case study participants explained changes to be made in their second drafts. Liu, for example, in response to her peer’s feedback, “What you wrote in this composition is about **literature**. You may either delete the comparison [between literature and text-message] or add some content of **text-message**” (Liu’s PF form), decided to delete the segment “rather than reading text-message” in her first draft. She explained as follows,

### Excerpt 5.7

(Liu, PF-triggered revision episode 18)

I will delete this part [“rather than reading text-message” in the sentence “Just like what the mother in the picture says, we should pay more attention to reading literature rather than reading text-message”], because it was not mentioned in this paragraph.

In Excerpts 5.6 and 5.7, the case study participants justified by explaining what they had written in their drafts. In addition, the participants also relied on their L2 knowledge when providing a justification. When Zhang responded to the feedback on the lack of link verb in the sentence segment “when people using phones, they can easily be attracted by another thing”, she provided a justification as follows,

### Excerpt 5.8

(Zhang, PF-triggered revision episode 7)

Since I had “**people**” as the subject in the time clause, I cannot omit the predicate. When I wrote [my first draft], I intended to use the pattern “**when doing**”. How could I use “**people**”, but forget to add the predicate? I can either omit the subject and the predicate altogether or keep the subject and predicate.

In this example, Zhang used her grammatical knowledge of time clauses to explain the problem in her first draft. Her metalinguistic explanations undoubtedly demonstrated her understanding of the problem as well as its solutions.

In the following excerpt, Sun drew on her L2 writing knowledge in explaining her modifications to the topic sentence of a body paragraph.

### **Excerpt 5.9**

(Sun, PF-triggered revision episode 3)

The [topic] sentence [in the second paragraph] should be “**First of all, reading literature lets us have a good command of words and structures.**”

I should put “words” before “structures” because the supporting sentence about “words” were presented first in this paragraph. The information in the topic sentence should correspond to that in the supporting sentences.

In this excerpt, Sun explained why she should put “words” before “structures” in her revised topic sentence. The terms “topic sentence” and “supporting sentences” in her explanation clearly indicate that she drew on her L2 writing knowledge to explain the change to be made in her second draft.

#### **5.3.1.5 Resourcing**

Other than *referring to PF* and *generating changes/no change*, *resourcing* was the most frequently employed cognitive operation in revisions triggered by peer feedback (see Table 5.4). As evident in their think-aloud protocols, the case study participants only consulted online and/or paperback dictionaries, to find an L2 word to use in their second drafts or check the meaning or use of an L2 word. The following excerpt illustrated how participants used online dictionaries to obtain L2 words.

### **Excerpt 5.10**

(Liu, PF-triggered revision episode 3)

[To search for a word to replace “considerable” in the sentence “The mother’s view is considerable and wise”] I am using the Kingsoft software [an online dictionary]. [Reading the word provided by the online dictionary] The English expression for “有意义的 [meaning significant]” is “**significant**”. I am searching for another word, “有参考价值的 [meaning valuable]”. [Reading the word provided by the online dictionary] I can use “**valuable**”.

Excerpt 5.10 illustrates how the case study participant utilized the Chinese-English translation function of an online dictionary to look for an L2 word to use in their second drafts.

Excerpts 5.11 to 5.13 show how the participants used online dictionaries to check the meaning of, the use of, and the spelling of words respectively.

### **Excerpt 5.11**

(Shen, PF-triggered revision episode 9)

What is the meaning of “**perseverance**”? It means “毅力, 韧性, 不屈不挠的精神 [meaning determination and the quality of not giving up]”.

### Excerpt 5.12

(Zhang, PF-triggered revision episode 2)

I was not sure about the collocation for “**benefit**”. Which proposition should I use after it, “**to**” or “**for**”? I will look up “**benefit**” in the online dictionary. It is “**benefit to** [*sic*]”. “**Benefit**” is a noun, and “**beneficial**” is an adjective.

### Excerpt 5.13

(Liu, PF-triggered revision episode 11)

[Using the online dictionary] I will check the spelling of “**horizon**”.

In addition to the two types of dictionary use, the analysis reveals that the six case study participants differed significantly in their frequency of dictionary use. Of the six case study participants, Wu used dictionaries least frequently (once in his total of 10 PF-triggered revision episodes). It appears that his trust in his peer meant that, in most revision episodes, he revised as his peer suggested rather than using dictionaries. The following excerpt is the only instance when he looked up a word, “aware”, to confirm the feedback point asking him to revise “Some people may have awared [*sic*] that...” in his first draft into “Some people may have been awared [*sic*] that...”.

### Excerpt 5.14

(Wu, PF-triggered revision episode 3)

I will look up “**aware**” in an online dictionary. [In the online dictionary] there are “**As you are aware**”, “**As far as I am aware**”, “**I don’t think people are really aware**” and “**be aware of**”.

Paying no attention to the part of speech of “aware”, Wu directly proceeded to an example sentence and its collocation. Although he finally came up with the right form, “Some people have been aware that” in his second draft, based on an inference from the examples in the dictionary, the process through which he worked out the right form suggested that he still did not understand his revision.

Liu, the case study participant who made the most extensive use of online dictionaries in revision (27 times for her total of 22 PF-triggered revision episodes), differed from the other three case study participants in the use of dictionaries. To some extent, Liu employed this cognitive operation excessively and ineffectively, in that, of the 27 instances of consulting dictionaries, there were four instances when she worked on spelling or morphological changes of specific words. For example, she checked the spelling of “technology” and the past tense for “put”. Furthermore, when consulting dictionaries, she focused on the Chinese meaning of a word more than on its use. When she obtained an L2 word by resorting to the Chinese-English function of online dictionaries, she directly used it in her second draft. In some other instances, she skipped over its collocation and example sentences and only read aloud its Chinese meaning.

The other three case study participants, however, demonstrated a tendency to use online dictionaries in conjunction with other cognitive operations. In most instances, they first retrieved an L2 word or phrase from their long-term memory, consulted online dictionaries for information that they needed, and then used it in their second drafts. When they used online dictionaries to obtain an L2 word, they looked it up in online dictionaries again to check for its use.

### 5.3.1.6 *Translating*

*Translating* is a cognitive operation in which the case study participants used their first language, Chinese, as a base for understanding or producing the second language. Seen from the think-aloud protocols, students translated what they wrote in English (L2) into Chinese (L1), and also translated ideas they generated in L1 into L2. The following excerpt taken from one revision episode illustrates the two functions of *translating*. It occurred when Zhang was adding some content as suggested by her peer,

#### **Excerpt 5.15**

(Zhang, PF-triggered revision episode 9)

Now I will translate this sentence [“Usually people tend to scan the news or text others to contact by [*sic*] their phones, which cannot leave a deep impression on themselves. But reading literature can give people more time to think about what they can learn from the author.”] into Chinese, “人们在手机上浏览新闻或者编辑短信给其他人的时候，这种方式不足以给自己留下一个很深刻的印象，但是在读文学作品的时候就可以给人们更多的时间去思考我们能从作者身上学到什么”...[In the revised sentence] I should add “从而对自己产生一个影响 [meaning “have an impact on themselves”].” This part should be “**and has a profound impact on themselves.**”

In the first half of this excerpt, Zhang translated the two sentences in her first draft directly into Chinese, even without reading them aloud. At this stage, *translating* seemed to help Zhang make the meaning of the text, fulfilling a similar function as

*referring to text.* Towards the end of this excerpt, Zhang said what she intended to add first in Chinese and then translated it into English literally. At this stage, *translating* helped Zhang generate a new text segment.

There were differences, however, in how the case study participants employed translation when generating a new text segment. Sun, unlike Zhang who generated new text through literal translation, used L1 only to decide what to write.

### **Excerpt 5.16**

(Sun, PF-triggered revision episode 9)

How should I revise this part [the concluding paragraph]? “**To summarize**”, what should I write after it? [Planning her ideas in Chinese] I can summarize the benefits of reading literature to form a contrast with reading from the internet.

In Excerpt 5.16, Sun verbalized only the gist of what she intended to write in her concluding paragraph, that is, a summary of the benefits of reading literature followed by the need to read literature attentively. Her concluding paragraph in the second draft was an elaboration of what she planned in Chinese, as follows,

To summarize, they can learn how to write by finding how literature tells and they can discover how the world out of sight is going on by reading what literature conveys. Compared to the incomplete and shallow knowledge that the Internet gave [*sic*] us, literature is more of a wise teacher, in which case, I’m sure people are expected to cover literature with patience on a regular basis. (Sun, Draft 2 of Writing Task 6)



### 5.3.1.7 *Monitoring*

*Monitoring* is a cognitive operation in which the case study participants directed their attention to a specific segment within a text when they were implementing revisions triggered by peer feedback. They carried it out in two ways: checking the accuracy and/or appropriateness of changes or attending to other information related to their changes.

The following excerpt occurred when Liu was adding details to “other aspects” in the sentence “It’s clear that reading can open our eyes, increse [*sic*] our knowledge, make us more thoughtful and improve ourselves in many other aspects.” She reminded herself to check the spelling of “technology”.

#### **Excerpt 5.17**

(Liu, PF-triggered revision episode 13)

How do I spell “**technology**”? Did I spell it correctly?

Similarly, other participants checked the changes they made from other aspects. For instance, when Sun was writing the revised sentence “Besides, the unique and attractive structure that the literature adopts to express its thesis precisely is also free for them to learn”, she reminded herself to use the singular form “structure” and add “-s” to “adopt” (Sun, PF-triggered revision episode 7). When Zhang was adding “has a profound impact on themselves”, she reminded herself not to use “leave a deep impression”, which she had already used (Zhang, PF-triggered revision episode 5).

In other instances, students consciously examined the untreated text segments and made corresponding changes. For instance, when Shen decided to replace “text-message” [in

the sentence “The picture shows that many young people prefer to receive and read text-message rather than literature nowadays.”] with “fragmentary reading”, he reminded himself to make related changes in other parts of the sentence.

### **Excerpt 5.18**

(Shen, PF-triggered revision episode 2)

“**Receive**”? How can I revise “**Receive and read text message**”? I need to delete “**receive**” and use the adverb form of “**fragmentary**”.

Another example was from Sun. On receiving the feedback point that a personal pronoun in her writing was inappropriately used, Sun checked all pronouns she used in her first draft: “If I want to adopt a third person perspective, I should use third person pronouns from the first paragraph.” (Sun, PF-triggered revision episode 2). As she read aloud, she stopped every time she came upon a pronoun and pondered over whether a change was needed.

#### **5.3.1.8 *Evaluating***

*Evaluating* is a cognitive operation in which the case study participants checked a change against a standard (received feedback or their own judgement about its appropriateness). Specifically, the case study participants evaluated changes they had made as well as changes that they intended to make. The following excerpt illustrates the evaluation operation that Wu carried out after implementing a change.

### **Excerpt 5.19**

(Wu, PF-triggered revision episode 4)

Now I have made my second paragraph shorter, as my peer suggested.

Excerpt 5.19 occurred after Wu deleted the sentence “Only because phones are quickly to search, it can’t be the reason we keep using it.” The deletion was made in response to a feedback point asking him to shorten his second paragraph. By carrying out the evaluation operation, Wu appeared confident that the changes that he made addressed the feedback he had received.

The case study participants, however, did not conduct the evaluation operation only after they implemented changes. In some instances, they evaluated changes they intended to make. When they judged that the changes were appropriate, they implemented them; when they judged that the changes were inappropriate, they implemented alternative cognitive operations. In the following excerpt, Shen evaluated a change before implementing it.

### **Excerpt 5.20**

(Shen, PF-triggered revision episode 1)

Will the sentence sound repetitive with two “**is**”?

The evaluation occurred after Shen verbalized his intention of adding “is” [after “and” in the sentence, “In the picture, a teacher is giving a thick book to a student and asking him to read the book just like reading a long text-message”]. Since he judged that the change would make his sentence repetitive, he revised it by using present participle as

an adverbial modifier [“In the picture, a teacher is giving a thick book to a student, asking him to read the book just like reading a long text-message”]. Also of note in this example is that Shen made the evaluation according to his L2 intuition, a subjective measure. The change he verbalized initially was grammatically accurate, but it did not sound right to him. He therefore used his grammatical knowledge and revised in another way.

#### 5.3.1.9 *Generating changes/no change*

Similar to *referring to PF*, the cognitive operation of *generating changes/no change* was present in each revision episode. In practice, the case study participants verbalised changes they made in their second drafts, or stated their decision to retain their original writing, as can be seen in the following two excerpts respectively.

##### **Excerpt 5.21**

(Wu, PF-triggered revision episode 8)

I write “**What’s more, the best advantage of reading books is to broaden our horizons**” [in my second draft].

##### **Excerpt 5.22**

(Sun, PF-triggered revision episode 5)

I will keep what I have written in my first draft [in response to the peer feedback point asking her to combine two sentences].

Within the revision episodes in which the case study participants accepted the feedback points, not all changes were generated within one step. For instance, when Liu added a

sentence as a transition between two paragraphs, she first came up with an interrogative sentence, “So why is it so important?” which she evaluated as “*not well-written*” (Liu, PF-triggered revision episode 10). She then engaged in the cognitive operation of *translating* to generate the sentence pattern “Why is reading...?” and *resourcing* to generate the word “position”. For the segment “to such a necessary position”, Liu generated it after consulting an online dictionary to check the collocations of “put” and “position”.

### 5.3.2 Approaches

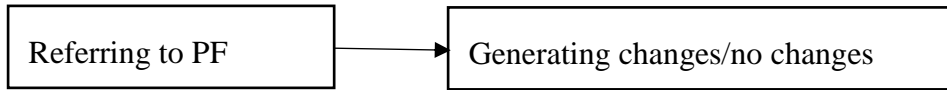
Based on the analysis of each PF-triggered revision episode for its cognitive operations, the present study identified two approaches of PF-triggered revisions: the direct approach and elaborated approach, as shown in Table 5.5.

When using a direct approach, the case study participants used two cognitive operations: *Referring to PF* and *generating changes/no changes*. The elaborated approach, however, included an extensive revision stage in which the case study participants deliberated over a received feedback point and corresponding changes through various cognitive operations. In the following sub-sections, each approach and factors influencing the case study participants’ use of it will be identified with excerpts from the think-aloud protocols.

Table 5.5 The direct approach and elaborated approach

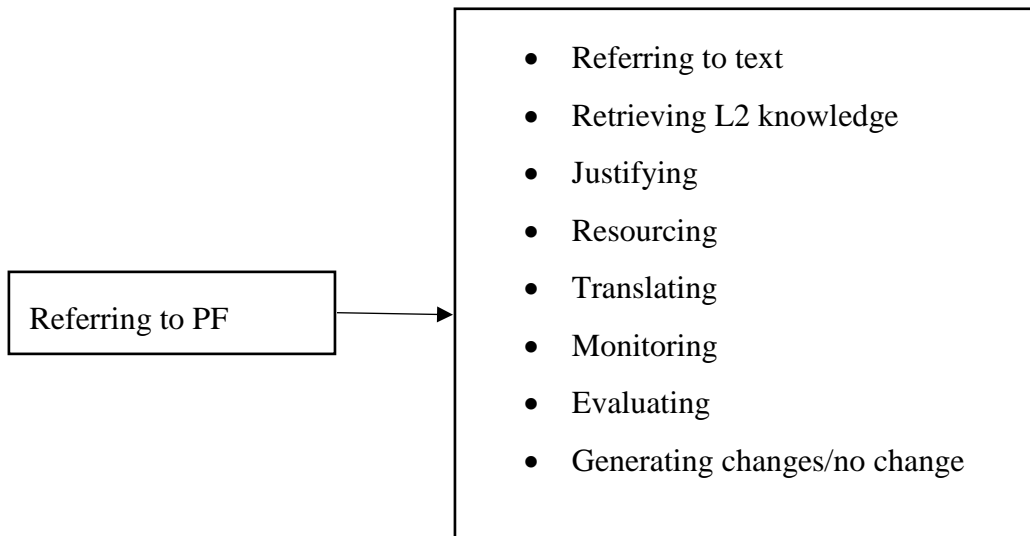
---

**Direct approach**



---

**Elaborated approach**



---

5.3.2.1 *The direct approach*

The direct approach, in which case study participants referred to feedback points they received and then generated changes or no change, was used predominately with feedback on language use, particularly mechanical errors and grammatical errors. The following example illustrates the direct approach.

**Excerpt 5.23**

(Liu, PF-triggered revision episode 5)

Here [The sentence segment “we should read books in a correct attitude”], my peer told me to use “**with**”, not “**in**”. <*Referring to PF*> I revised “**in**” into “**with**”. <*Generating changes*>

In Excerpt 5.23, Liu referred to the feedback point she received on her PF form, “with a(n)...attitude” and stated her decision to adopt it. With feedback on mechanical and grammatical errors, probably because revision solutions were included in feedback points and only simple changes were required, the case study participants had a strong tendency to use the direct approach.

The following revision episode from Wu, although similar to the above example, suggest that discussion was another factor leading to the employment of direct approach.

#### **Excerpt 5.24**

(Wu, PF-triggered revision episode 5)

Yes, my peer told me to use passive voice. <*Referring to PF*> [I will write]  
“**They can all be known from the books.**” <*Generating changes*>

(Discussion between Wu and his peer)

**Peer:** “**They can all know from the books**”; this is a wrong sentence. You should use passive forms.

**Wu:** I have forgotten what I wrote.

**Peer:** Passive form here.

**Wu:** I should have written “**They can all know from the books**”. Does the

sentence sound right? How can I say it in passive voice?

**Peer: They can all be known from the books.**

**Wu:** Yes, I get it.

Excerpt 5.24 occurred in response to a feedback point asking Wu to use passive form. Wu's peer, however, did not include the correct form in his feedback. Even though the required changes were not difficult, Wu obviously could not come up with the correct form on his own, as evidenced from his incorrect revision ("They can all know from the books.") and his uncertainty about passive forms throughout the discussion ("Does the sentence sound right?" and "How can I say it in passive voice?"). As his peer provided him with the correct form orally, and Wu indicated his agreement in the discussion, it was more likely that their discussion led to Wu's use of the direct approach.

In another revision episode, Sun rejected a feedback point on a transition between two sentences based on the discussions about it.

### **Excerpt 5.25**

(Sun, Draft 1 of Writing Task 6)

"Without literature, the words we use are confined to what we hear and see in daily life, which are usually common and boring. Nevertheless in literature, it wouldn't be hard for us to find words that convey message differently but accurately."

(Sun, received feedback on PF form)



The two sentences linked by “**nevertheless**” should be combined into one sentence.

(Sun, PF-triggered revision episode)

“**Nevertheless**”. <*Referring to PF*> No, I won’t revise it <*Generating no change*>.

(Discussion between Sun and her peer)

**Peer:** “Nevertheless” and “but” means the same thing.

**Sun:** Yes, but “nevertheless” is an adverb, and that’s the difference between them.

**Peer:** The two sentences are related.

**Sun:** Yes, they are related. The logical relation between the two sentences is contrast.

**Peer:** If they are related, you should not use two sentences. Why do you use “nevertheless”? Is it because the word is more appropriate or because the word looks advanced?

**Sun:** I think “nevertheless” is appropriate here. The former sentence is about the words we used in daily life and the latter mentioned the words we obtained from reading literature. The two sentences are opposite in logic.

**Peer:** Yes, but you separated the two sentences.

**Sun:** Otherwise the sentence will be very long, taking up four lines. Too long

a sentence exhausts your readers. You don't want the two sentences to be separated, do you?

**Peer:** I still suggest you revise “nevertheless” into “while” and combine the two sentences. That's my opinion.

The discussion between Sun and her peer was extensive: Sun explained what she wrote and why she used “nevertheless” in her first draft while her peer explained the problems he identified and the change he suggested. By clarifying “nevertheless” in terms of its part of speech and explaining its appropriateness in joining the two sentences expressing opposite ideas as well as the need to consider sentence length, Sun indicated her intention to retain her original expression. Since the decision to reject the feedback point was probably made during the discussion, Sun mentioned the feedback point and her decision to reject it only briefly in the revision episode.

### 5.3.2.2 *The elaborated approach*

When adopting the elaborated approach, the case study participants used various cognitive operations in the revision stage; they implemented changes in multiple steps having considered feedback they received. The case study participants used this approach in response to feedback on complex language problems as well as on content and organization.

The nature of changes required by feedback was a possible reason for the case study participants' use of the elaborated approach. The following example illustrates how Yang employed the elaborated approach in dealing with a feedback point concerning sentence structure.

#### **Excerpt 5.26**

(Yang, received feedback on PF form)

Your description about the picture is too long.

(Yang, PF-triggered revision episode 1)

My peer commented that my introduction paragraph was long and suggested that I make my description about the picture short. [Reading the corresponding sentences in his first draft] “From this picture, we can see that here’s a teacher handing a book to a little boy. **‘Just thinking [sic] of it as if you’re reading a long text-message’, That’s what the teacher saying [sic] to the little boy. As far as the picture is concerned, we can get the importance of reading literature from the little boy’s action that is to go to read literature.**” This part was indeed long. <Referring to PF> I should delete some repetitive expressions. “**As far as the picture is concerned** “can be deleted and some information in these sentences can be combined. <Generating changes> I can use, “**the action of**” [to revise segment “here’s a teacher handing a book to a little boy” into “the teacher’s action of handing a book to the little boy”]. <Retrieving L2 knowledge> I also need to use “**her words**” to replace “**That’s what the teacher saying**”. <Monitoring> The segment “**from the little boy’s action that is to go to read literature**” can be deleted. <Generating changes> So I will write “**From this picture, we can get the importance of reading literature from the teacher’s action of handing a book to the little boy and her words “Just thinking [sic] of it as if you’re reading a long text-message**”. <Generating changes> The revised sentence is long, but I use one sentence instead of three <Evaluating>.

This revision episode was in response to a feedback point asking Yang to shorten his description about a picture. In contrast to feedback on mechanical and grammatical errors, feedback on sentence structures obviously required complex changes. In this example, Yang employed six cognitive operations in four categories in the revision stage. After applying the feedback point back to the text, he used his grammatical knowledge about nominalizations, monitored other related text segments, generated changes in three steps and finally evaluated the revised sentence in his second draft. His revised sentence, “From this picture, we can get the importance of reading literature from the teacher’s action of handing a book to the little boy and her words ‘Just thinking [*sic*] of it as if you’re reading a long text-message””, was more concise in describing the picture (two lines shorter than his original description) and more effective in introducing the topic on the importance of reading literature.

Similarly, the case study participants used the elaborated approach to deal with feedback on content and organization. In the following revision episode, Zhang responded to a feedback point which asked her to enrich the content in the concluding paragraph.

### **Excerpt 5.27**

(Zhang, PF-triggered revision episode 10)

My concluding paragraph is too short with only one sentence, and my peer suggested that I add some content, such as the main points in my composition.

*<Referring to PF>* Yes, I can summarize the points in my three body paragraphs. I can make some changes to my thesis statement. [Reading some segments in the thesis statement] I used “**understand what we learn better**”, “**study more effectively**”, and “**has benefit to our health**”. *<Referring to*

*text*> I can use the noun form “**understanding**”. <*Retrieving L2 knowledge*>  
Which verb should I use together with it? [Reading from the online dictionary]  
I can use “**gain an understanding of**”. <*Resourcing*> How about “study  
more effectively”? I can use “**improve learning efficiency**”. <*Retrieving L2  
knowledge*> How about “**has benefit to our health**”? No, I cannot use  
“**benefit**”. In the previous sentence, I wrote “**benefit much from it**”.  
<*Monitoring*> I use “*do no harm to*”. <*Retrieving L2 knowledge*>. But I need  
to put this point in a separate sentence, because I used negative form for this  
point. <*Justifying*> I need a conjunction, maybe “**in addition**”. <*Retrieving  
L2 knowledge*> So I add “**Reading books can help us achieve a better  
understanding about our learning and improve our knowledge. In  
addition, it does no harm to our health**”. <*Generating changes*>

In this revision episode, Zhang added the three points mentioned in her body paragraphs to strengthen her concluding paragraph. Since addition of content requires generation of new ideas and consideration of language forms, Zhang revised her draft extensively and implemented nine cognitive operations. When generating ideas for the additional sentence, Zhang referred first to the thesis statement: “understand what we learn better”, “study more effectively”, and “has benefit to our health”, and then considered language forms for her new sentences. She drew on her vocabulary knowledge (for “understanding”, “improve learning efficiency”, “do no harm to” and “in addition”) and consulted online dictionaries (for the collocation “gain an understanding”), while at the same time, reminding herself to avoid word repetition and providing justifications for starting the third point with a new sentence. After carrying out these cognitive operations, Zhang generated two sentences for her second draft, “Reading books can help us achieve a better understanding about our learning and improve our knowledge” and “In addition,

it does no harm to our health”, which improved her concluding paragraph in content and in length.

This revision episode contains two aspects of the case study participants’ use of the elaborated approach in response to feedback on content and organization. First, the case study participants typically implemented more cognitive operations in response to feedback points on content and on organization, than on language use. Second, they showed a strong tendency to monitor the accuracy or appropriateness of their changes, because feedback on content and organization also requires attention to language forms.

Another factor contributing to the case study participants’ employment of the elaborated approach was a critical attitude towards the received feedback. Although feedback points on both spelling errors and grammatical errors and feedback points on lexical errors included suggested changes, the case study participants dealt with them differently. In most instances, they employed an elaborated approach in response to feedback points on lexical errors, confirming the validity of the feedback points before incorporating them in their second drafts, as exemplified in the following revision episode.

### **Excerpt 5.28**

(Shen, received feedback on PF form)

“**Persist**” is always used as an intransitive verb, meaning “continue to do something, although it is difficult” and its collocation is “**persist in**”. You can revise it [in the sentence segment “reading literature improves our ability to persist”] into “**reading literature helps develop our perseverance**”.

(Shen, PF-triggered revision episode 10)

“**Persist**” is always used as an intransitive verb, meaning “continue to do something, although it is difficult” and its collocation is “**persist in**”. My peer suggested that I use “**reading literature helps develop our perseverance**”.

<*Referring to PF*> I need to look up “**perseverance**” in the dictionary. “**Perseverance**” means determination to achieve something in spite of difficulties. I can use it here. <*Resourcing*> How should I use “**help**”, “**help do**” or “**help to do**”? [Reading from the dictionary] Both of them are right.

<*Resourcing*> I will write “**reading literature helps develop our perseverance**”. <*Generating changes*>

This revision episode occurred in response to a feedback point asking Shen to replace “persist” with “develop perseverance”. The feedback point was very specific in that the reviewer explained the problem (the meaning and collocation of “persist”) and suggested changes (“help develop our perseverance”). Rather than directly incorporating the suggested changes, Shen engaged in an extended stage in which he consulted online dictionaries twice, first for the meaning of “perseverance” and again for the collocation of “help”. Having confirmed the feedback point, he revised as suggested by his peer.

The following example from Yang illustrates how a critical attitude and the adoption of an elaborated approach helped him generate more appropriate changes than what was suggested by his peer.

### **Excerpt 5.29**

(Yang, received feedback on PF form)

“**Learn morality**” [in the sentence “Due to the thinkings [*sic*] we do while reading, we learn lots of morality which influences us all the time during our reading”] should be revised into “**raise the moral level**”.

(Yang, PF-triggered revision episode 6)

“**Morality**” is problematic. My peer told me to use “**raise the moral level**”.  
<*Referring to PF*> In Chinese, the phrase means “道德水平”. That is the meaning I wanted to express. <*Translating*> What is the English equivalent for “道德水平”? [Reading from the online dictionary] There are “**moral level**” and “**moral standard**”. <*Resourcing*> Can I say “**moral level**”? It sounds a little weird. Let me check the use of “**moral**”. [Reading from the online dictionary] There is “**moral standards/values/principles**”, but not “**moral level**”. <*Resourcing*> “**Raise**” can be used together with “**standards**”. <*Retrieving L2 knowledge*> OK, I will use “**raise the moral standards**”.  
<*Generating changes*>

This revision episode occurred when Yang was dealing with feedback on his problematic use of “morality”. As in any other instances, Yang first referred to the feedback point he received by mentioning the problem as well as the suggested changes. Probably because Yang was unfamiliar with the expression “raise the moral level”, he engaged in a revision stage in which he translated his peer’s suggestion into Chinese and used online dictionaries to obtain two English equivalents. After he had confirmed the expression “moral standards”, he used his vocabulary knowledge and decided to use “raise the moral standards”, a more idiomatic expression than suggested by his peer, in his second draft.



#### **5.4 Chapter summary**

This chapter, based on the PF-triggered revision episodes from the think-aloud protocols, presented how the case study participants utilized the nine cognitive operations (*referring to PF, referring to text, retrieving L2 knowledge, justifying, resourcing, translating, monitoring, evaluating and generating changes/no change*) and the two approaches (the direct approach and the elaborated approach) when processing peer feedback points in revision. Further analysis showed that the case study participants tended to use the direct approach when they dealt with feedback points on simple language problems (spelling errors and grammatical errors). In response to feedback points on complex language problems (sentence structure and vocabulary) and on content and organization, the case study participants demonstrated a tendency to use the elaborated approach. The nature of changes required by feedback was identified as a common factor influencing their use of the two approaches. Discussion was a factor that led to their use of the direct approach while a critical attitude was a factor that contributed to their use of the elaborated approach.

## Chapter 6 Student Perceptions of Peer Feedback

### 6.1 Chapter overview

This chapter reports the treatment group participants' perceptions of the intervention, addressing the following question and its sub-questions.

RQ 4 How do Chinese EFL learners perceive the effectiveness of peer feedback intervention?

*How do Chinese EFL learners perceive their changes in writing performance?*

*How do Chinese EFL learners perceive the usefulness of specific components in the peer feedback intervention?*

Data were collected from the open-ended questionnaires completed by the 35 treatment group participants and the semi-structured interviews with the six case study participants in Week 16.

This chapter is divided into four sections: a chapter overview, two sections reporting the treatment group participants' perceived changes in the six measures of writing performance and their perceptions of the usefulness of specific components in the intervention, and a chapter summary. In each of the two sections, the data on the treatment group participants' attitudes towards each statement and their explanations are presented separately. A priority is given to their explanations, with the topics identified from their explanations summarized and sample quotes showing their representative perspective provided. Excerpts from interviews with the six case study participants will also be quoted when relevant.

## 6.2 Perceived changes in writing performance

Three items in the questionnaire elicited the treatment group participants' perceptions of changes in their writing performance as targeted in the quasi-experimental study.

Table 6.1 presents the three questionnaire items and the treatment group participants' extent of agreement with the statements. As seen from the table, the majority of the treatment group participants claimed that they improved in each of the three aspects of writing performance, with a very high percentage agreeing that they improved in content and in organization of their compositions and no participant disagreeing with either of the two statements. Although several treatment group participants were neutral or disagreed with the statement on language use, the majority acknowledged their improvements. The following subsections present the analysis of the treatment group participants' explanations to each questionnaire item.

Table 6.1 Treatment group participants' perceptions of their improvements in writing performance

Questionnaire Items	Participants (N = 35)		
	Agreed	Neutral	Disagreed
By participating in peer feedback activities all through the semester, I have improved in the content of my composition.	33 (94.3%)	2 (5.7%)	0
By participating in peer feedback activities all through the semester, I have improved in the organization of my composition.	33 (94.3%)	2 (5.7%)	0
By participating in peer feedback activities all through the semester, I have improved in the language use of my composition.	31 (88.5%)	3 (8.5%)	1 (2.8%)

### 6.2.1 Content and organization

Topics identified in the comments on content and organization, 33 and 35 respectively, are presented in Table 6.2 and Table 6.3. The treatment group participants were generally confident about their improvements in content and organization of writing, with only two participants identifying no change in organization of their compositions.

Table 6.2 Students' perceived changes in content

<b>Comments</b>	<b>Frequencies</b>
Improvement in relevance: Fulfilling task requirement	13
Improvement in relevance: Avoiding irrelevant information	11
Improvement in idea development	9
Total	33

Table 6.3 Student' perceived changes in organization

<b>Comments</b>	<b>Frequencies</b>
Improvement in text structure	19
Improvement in paragraph structure	10
Improvement in cohesion	5
No change in text structure	2
Total	35

Typically, the treatment group participants, in comparing their compositions for the first few writing tasks with those for the last few writing tasks, identified areas they felt improved.

**Excerpt 6.1**

After participating in peer feedback activities several times, I could write a thesis statement as required by the directions (Participant #3).

**Excerpt 6.2**

I have learned different methods of idea elaboration, such as exemplification and comparison. Before I participated in this project, I only elaborated ideas by explaining. (Participant #10)

**Excerpt 6.3**

When I compared the body paragraphs I wrote for the first three writing tasks and those for the remaining ones, I noticed the differences. The body paragraphs for the first three writing tasks were very loose. In some paragraphs, I did not include topic sentences and in others, I did not write topic sentences until the end of a paragraph. (Participant #7)

**Excerpt 6.4**

Before my participation in this project, I never thought about structures. I wrote whatever I thought of and put the ideas randomly in my composition. But now I know I should introduce my thesis statement in the introductory

paragraph, present one argument in each body paragraph, and summarize all my arguments in the concluding paragraph. (Participant #24)

In the above examples, the participants were specific about the areas they felt improved: Participant 3 mentioned the pertinence of his thesis statement to the directions of writing tasks; Participant 10 identified strategies for idea development; Participant 7 referred to the paragraph structure, and Participant 24 mentioned text structure. Since their explanations were metalinguistic and contextualized, it is probably they had gained some knowledge about content and organization in writing and utilized it in their compositions.

As well as identifying areas in which they made progress, the participants mentioned the elements within the intervention which contributed to their improvements, as illustrated below.

#### **Excerpt 6.5**

Some feedback points from my peers were very useful, such as those about consistency of pronouns and relevance of information. (Participant #4)

#### **Excerpt 6.6**

When I was reviewing my peers' compositions, I was learning from their strengths and weaknesses. For instance, after evaluating the relations between the arguments and the thesis statement in their compositions, I examined my own arguments. Sometimes, I even incorporated their arguments in my writing. (Participant #35)

### **Excerpt 6.7**

For the first few writing tasks, writing a concluding paragraph was really a headache for me, as I could come up with only one sentence. Around the middle of the semester, I had a very meaningful discussion with my peer on how to write the concluding paragraph. She explained her concluding paragraph to me and introduced some other ways. Now, I can write an effective concluding paragraph. (Participant #40)

In the above excerpts, the three participants identified the areas they improved, and more importantly, linked their improvements to specific components in the intervention. They paid attention to content and organization problems that had been identified by peers, drew on their peers' compositions to strengthen their own writing and to avoid similar problems, and used discussions to solve problems they had in content and organization.

Similarly, when asked to name one area (content, organization, or language use) that they felt improved most significantly, five case study participants identified organization and one identified content. The following example from Yang provides evidence of his improvement in text structure.

### **Excerpt 6.8**

The most significant change in my writing is paragraphing. For the pre-test and the first two writing tasks, I used the three-paragraph model, an introductory paragraph, a body paragraph including three points, and a concluding paragraph. When the teacher presented the five-paragraph structure [in Writing Task 3] in class, I used it in my revised draft. Probably because I felt it difficult to come up with supporting ideas, I went back to the

three-paragraph structure in the subsequent two writing tasks. In Writing Task 6, I realized the need to paragraph when I was reading my peer's composition. Her content was not well-written, but I had no difficulty understanding the ideas because of her text structure. Within the same task, I received her feedback on this problem. For my second draft of Writing Task 6, I revised my long body paragraph into three separate ones. In the post-test, because of the time constraint, I included two arguments in my composition and wrote two body paragraphs. (Yang, interview)

In this excerpt, Yang first explained that the change in text structure was not easy to achieve. Although the teacher's instruction guided Yang's use of five-paragraph structure in his revised draft, he regressed to the three-paragraph structure in two writing tasks. Secondly, reviewing his peer's composition and receiving feedback together led to Yang's change in text structure. His grasp of text structure can be seen from his flexible use of two body paragraphs in the post-test.

Text structure was most frequently mentioned by the treatment group participants (see Table 6.3). Some, similar to Yang, started to make changes in response to feedback from their peers, while others attributed their change to their teachers' explanation of text structure in the focused training session [in Writing Task 3], as seen in the following excerpt.

### **Excerpt 6.9**

I began to organize my arguments into separate paragraphs after the teacher's instruction. Probably because the paragraphing problem was common among us, the teacher addressed it in class. I remembered that she asked us to



compare the two drafts written by one of my classmates, one with a long body paragraph and the other with three separate ones. My first impression was that the latter one looked neat and clear and the topic sentence within each paragraph stood out. She then explained the five-paragraph pattern in expository and argumentative writing. All my subsequent compositions followed the five-paragraph structure. (Participant #16)

The two students, who assumed that they made no progress in organization, also mentioned the text structure. Their explanations, however, are different.

#### **Excerpt 6.10**

There was no obvious change in text structure. I kept using the the three-paragraph structure although some of my peers suggested that I include more body paragraphs. Different people have different writing styles and distributing the arguments to separate paragraphs is just one option. (Participant #20)

#### **Excerpt 6.11**

I do not have many problems with organization in writing. Since I was introduced to the five-paragraph structure when I was preparing for other tests, I used it all through the semester. (Participant #29)

Seen from Excerpt 6.10, Participant 20 noticed the differences in text structure without understanding why paragraphing was necessary. As most Chinese EFL students have been taught the three-paragraph structure in Chinese writing and English writing for many years, it is possible that her previous L1 and L2 writing experiences still had an

influence. For Participant 35, it is more likely that she felt already proficient in the organization of writing and saw no room for further improvement. These two excerpts, together with Excerpts 6.4, 6.8 and 6.9 and the frequencies reported in Table 6.3, suggest that students prioritized text structure when evaluating the impact of the intervention on their writing performance.

### 6.2.2 Language use

In response to the statement on language use, the treatment group participants provided 60 comments. This item had the highest response rate in the questionnaire, and their explanations referred to five topics, as presented in Table 6.4.

Table 6.4 Students' perceived changes in language use

<b>Comments</b>	<b>Frequencies</b>
Fewer errors in writing	27
Improved L2 knowledge: grammatical structures /new words	16
Enhanced attention to language problems	9
Recurrence of language problems	4
Avoidance of complex structures in writing	4
Total	60

As seen in Table 6.4, more than half of the treatment group participants identified a decrease in errors, particularly grammatical errors which were the most direct and observable evidence of improvement in accuracy. They mentioned their performance in overall accuracy as well as in specific errors.

**Excerpt 6.12**

I totally agreed with this statement. Within one writing task, there must be fewer grammatical mistakes in my revised draft because I made revisions. Over the semester, the number of grammatical mistakes in my writing has also decreased. Now, the fewer comments on grammar from my peers might suggest I have fewer grammatical mistakes. (Participant #7)

**Excerpt 6.13**

I have fewer mistakes in my writing, especially those concerning singular and plural forms and agreement between the subject and the predicate. (Participant #16)

**Excerpt 6.14**

Now, there are fewer “the” in my compositions. In the last few writing tasks, whenever I wanted to use “the”, I recalled my peers’ comments and gave it a second thought. (Participant #5)

In Excerpt 6.12, Participant 7 compared the number of grammatical errors in his two drafts of a writing task as well as in his compositions for different writing tasks. When evaluating progress in language use throughout the semester, the participant used the number of feedback points on grammar received for each writing task as a reference point: Fewer feedback points on grammar would indicate improved performance. In Excerpt 6.13 and 6.14, the two participants described specific errors, such as singular/plural form, agreement between sentence segments, and use of the definite article “the”, which were very common among EFL writers.

In addition, nearly half of the treatment group participants mentioned the value of the peer feedback activities in strengthening their L2 knowledge. Specifically, they mentioned their improved knowledge about some grammatical structures and their expanded vocabulary.

**Excerpt 6.15**

Some of my peers were very responsible. They corrected my errors as well as explained corresponding rules. Their explanations helped me understand the rules. (Participant #38)

**Excerpt 6.16**

When I was reviewing my peers' compositions, I sometimes needed to consult reference books. Such experience was also helpful for me. The next time I used these rules, I used them correctly. With more and more grammatical knowledge activated and strengthened, I have made fewer grammatical errors in my writing. (Participant #20)

**Excerpt 6.17**

The comments from my peers helped me accumulate words. They suggested words that I could use in my revised drafts. For some words, I studied their meaning and use by consulting dictionaries, and tried using them in my writing. (Participant #16)

In Excerpt 6.15, Participant 38 acknowledged that feedback with explanations helped him understand grammatical rules. In Excerpt 6.16, Participant 20 identified her use of reference materials when providing feedback for her peers. By consulting the reference

materials, she generated a reliable feedback point for her peer as well as clarified her own understanding of a specific grammatical rule. In Excerpt 6.17, Participant 16 claimed that receiving feedback on vocabulary and consulting dictionaries helped expand his own vocabulary. These three excerpts suggest that the peer feedback intervention provided students with opportunities to expand their L2 knowledge, a key resource for their writing.

### **Excerpt 6.18**

I know more words from participating in peer feedback activities, but I have difficulty using them in my writing. (Participant #5)

Excerpt 6.18 was one of the four comments in which the participants acknowledged they had difficulty in using the words that they learned through peer feedback activities in their writing. Similarly, in the interview, Liu admitted to experiencing difficulty using the newly-acquired words in her writing.

### **Excerpt 6.19**

**Liu:** Some of my peers used a lot of good expressions. Some of them conveyed the exact meaning I intended to express.

**Researcher:** Can you provide an example to illustrate?

**Liu:** Yes, [I remembered the phrase] “开阔眼界” [meaning “broaden one’s horizons”].

**Researcher:** Do you remember how to say it in English?

**Liu:** [Pausing for while] Sorry, I can’t.

**Researcher:** I will give you some hint, “horizon”. Which verb do you use together with it?

**Liu:** Yes, the phrase includes “horizon”. Sorry, I can’t remember it. For most of the new expressions, I only used them in my revised draft. When it came to a new topic, I still used the expressions I was familiar with. I can hardly think of the newly-acquired ones.

This excerpt occurred when Liu was asked to illustrate a word or phrase which she learned from his peers. Only knowing the Chinese meaning, Liu failed to recall the phrase even when she was given “horizon” as the hint and was unable to use it in her writing as she mentioned at the end of the excerpt. If students appear to know some words, but have difficulty using them in their writing, it is unlikely that their writing vocabulary has improved.

Some participants, however, referred to their increased attention to language problems when writing, as shown in the following excerpts.

**Excerpt 6.20**

Now I understand what grammatical categories I had problems with. For several writing tasks, my peers pointed out my spelling errors and my inconsistent use of tenses. For the last few writing tasks, I was very sensitive to these problems when writing my first drafts. Before turning them in for peer feedback, I read them through to check for these two types of errors. Such errors therefore decreased in my writing. (Participant #27)

**Excerpt 6.21**

After two of my peers told me to avoid word repetition in my writing, I constantly reminded myself to diversify my words when writing. Sometimes I used synonyms to avoid word repetition. (Participant #34)

**Excerpt 6.22**

I totally agree with the statement. Before I participated in this research, I used only attributive clause and object clause in my writing. Now, I could use non-finite verb forms, other types of noun clauses as well as compound sentences linked by conjunctions. Another important change in my writing was the use of simple sentences. I sometimes break my long sentences into several simple sentences to convey my intended meaning more clearly. (Participant #17)

In the above excerpts, the participants' enhanced attention to language was evident in their efforts to self-monitor the composing and self-editing processes. In Excerpt 6.20, Participant 27 mentioned her attention to spelling and tense errors, which were repeatedly pointed out by her peers. In Excerpt 6.21, Participant 34 referred to her attention to word repetition as well strategies to avoid the problem. In Excerpt 6.22, Participant 17 demonstrated an awareness of the need to use different grammatical structures instead of a few structures repetitively, as well as his use of short and simple sentences in writing. As most Chinese EFL students believe that using complex structures could achieve high scores in tests, some participants' realization that using simple short sentences can also be a language strategy in writing is a development.

However, eight participants commented on their lack progress in language use in writing, which they attributed to the recurring nature of errors and their avoidance of complex sentence structures.

### **Excerpt 6.23**

There are so many language problems in writing. Some problems were solved but new problems kept emerging. (Participant #12)

### **Excerpt 6.24**

I only write simple sentences in my compositions. Because of my low proficiency in English, I never take the risk of using complex sentence structures. (Participant #23)

Obviously, neither factor was pertinent to the peer feedback practice. These two excerpts, when examined together with those previously mentioned in which the participants identified their decrease in grammatical errors, suggest that some participants did not benefit from the intervention as much as their peers.

### **6.3 Perceived usefulness of specific components in the intervention**

Six questionnaire items concern the students' perceived usefulness of the components in the peer feedback intervention. As seen in Table 6.5, the majority of the treatment group participants perceived each component in the intervention to be useful.



Table 6.5 Treatment group participants' perceptions of specific components in the intervention

Questionnaire Item	Participants (N = 35)		
	Agreed	Neutral	Disagreed
I found the general training session on peer feedback implemented at the beginning of this semester useful.	33 (94.2%)	2 (5.7%)	0
I found the six focused training sessions on peer feedback implemented within each writing task useful.	34 (97.6%)	1 (97.1%)	0
I benefitted from reviewing my peers' compositions throughout the semester.	32 (91.4%)	3 (8.6%)	0
I benefitted from receiving feedback from my peers throughout the semester.	29 (82.8%)	5 (14.3%)	1 (2.9%)
I found the paired discussion session within each writing task useful.	28 (80%)	4 (11.4%)	3 (8.6%)
I found using the PF forms throughout the semester useful.	22 (62.9%)	5 (14.3%)	8 (22.9%)

### 6.3.1 General training and focused training

As described in 3.5.2, the present study implemented peer feedback training in two forms, a general training session and six focused training sessions. The general training session for the treatment group participants, prior to the intervention, included the essential concepts of peer feedback. The six training sessions, each focusing on one feedback or revision skill, took place alongside the six writing tasks. In response to the questionnaire items about the general training session and the focused training sessions, 33 and 34 comments were collected, respectively. Table 6.6 presents the topics identified from the participants' responses to the general training session.

Table 6.6 Students' perceptions of the general training session

<b>Comments</b>	<b>Frequencies</b>
Developing knowledge/skills for peer feedback activities	24
Fostering self-monitoring/evaluation ability in writing	9
Total	33

The topics most frequently recalled from the general training was development of knowledge and skills for peer feedback activities. Although there was a time interval of 15 weeks between the general training session and the questionnaire administration, the majority of the participants still recalled what they gained from the general training.

**Excerpt 6.25**

Since I had no experience in peer feedback, the general training helped me understand steps in peer feedback activities. I have known what I was expected to do at each step. (Participant #8)

**Excerpt 6.26**

I have known what I should focus on when giving feedback. I should focus more on ideas and organization than on vocabulary and grammar. I also know the elements I needed to include in a piece of feedback. Introduction to these concepts was useful, as I kept using them all through this semester. (Participant #17)

**Excerpt 6.27**

The teacher's explanations about the desirable attitude was useful. By being sincere and cooperative, we could avoid frictions and make the best of peer feedback activities. (Participant #22)

**Excerpt 6.28**

I used to have doubts about my own language proficiency and that of my peers'. The comparison between the comments we generated for the model composition and those my teacher generated was helpful. (Participant #16)

In the above excerpts, the participants mentioned the procedures, foci, stances and attitudes in peer feedback activities, and their confidence in participating in peer feedback activities, as "all factors contributing to effective peer revision" (McGroarty & Zhu, 1997, p. 19). The majority of the treatment group participants had no prior experience of peer feedback and so the general training session prepared them with "hands-on experience for the upcoming activities" (Participant #14), without which, peer feedback activities would be "chaotic and confusing" (Participant #30).

Surprisingly but encouragingly, some participants mentioned their self-monitoring and self-evaluation behaviours in writing due to the general training.

**Excerpt 6.29**

The general training was useful. I knew what to focus on when I am writing. Some terms the teacher introduced, such as thesis statement and topic sentences, came to my mind when I was writing. (Participant #25)

### Excerpt 6.30

After the training session, I started to use the procedures and strategies in my own writing and revision. When I finished my first draft, I read it through to detect possible errors. When I finished my second draft, I checked whether my revisions were all grammatical ones. If that was the case, I could read my draft again to check for possible problems in content and organization.

(Participant #9)

In the above two excerpts, it is evident that the participants responded to feedback focus and applied the reviewing procedures in their own composing and revising processes. By using their own compositions to rehearse what they learned from the training, they were able to adapt to the upcoming peer feedback tasks. More importantly, they demonstrated self-monitoring and self-evaluation behaviours, two aspects of self-regulation, which have potential for improving performance in writing.

When recalling aspects of focused training sessions, the treatment group participants identified the specific focus of each session and its contribution to revision (see Table 6.7).

Table 6.7 Students' perceptions of focused training sessions

Comments	Frequencies
Facilitating the revision	26
Appreciation of the specific focus in each session	8
Total	34

The following excerpts illustrate the influence of focused training sessions on revision.

**Excerpt 6.31**

The teacher analyzed the problems and offered several revision suggestions. She offered clearer explanations than my peers. Her explanations helped me make revisions. (Participant #4)

**Excerpt 6.32**

My teachers exemplified some problems that my peers did not detect in my writing. (Participant #40)

Participant 4 felt the teacher's explanations in the focused training session gave her a better understanding of some feedback points. For Participant 40, the focused training helped him identify problems which his peers did not detect. These two excerpts suggest that the focused training session supplemented or complemented peer feedback and supported their revisions.

Some participants, referring to their role as a feedback giver, explained how they quoted the teacher's explanations to support or strengthen their feedback, which facilitated their peers' use of it in revision.

**Excerpt 6.33**

The focused peer feedback training was useful, especially my teacher's explanations of some problems. When I was reviewing my peers' compositions, I did not comment on problems that I felt uncertain about. Luckily, similar problems were mentioned by my teacher in class. In the

discussion, I explained some problems as my teacher did in class. (Participant #28)

In Excerpt 6.33, the participant reported how the teacher's explanations helped him generate some feedback points which he had avoided because of uncertainty. Students, still EFL learners, could feel incapable of explaining some problems and therefore refrain from commenting on them at times. Once such difficulties were addressed by the teacher in the focused training sessions, students understood the problems better and felt able to provide reliable feedback and help their peers revise.

Some treatment group participants also indicated their appreciation of the specific focus within each training session. With a single focus in each session, they felt related topics were dealt with substantially. In the following excerpt, the participant recalled the two training sessions she felt most useful.

#### **Excerpt 6.34**

I found the training on grammar and on vocabulary extremely useful. For grammar, the teacher presented a long list of typical grammatical mistakes selected from our compositions and we together revised them one by one. For vocabulary, the teacher introduced a variety of resources and taught us how to use them. (Participant #31)

#### **6.3.2 Giving feedback and receiving feedback**

Table 6.8 and Table 6.9 present the topics identified from the treatment group participants' comments for the two questionnaire items on giving feedback and receiving feedback; 32 comments and 33 comments were collected respectively. Their

comments suggest the participants were cognizant of the value of giving and receiving feedback.

Table 6.8 Students' perceptions of giving feedback

<b>Comments</b>	<b>Frequencies</b>
Learning new ideas/specific skills	18
Training reader awareness	11
Providing motivational impetus	3
Total	32

Table 6.9 Students' perceptions of receiving feedback

<b>Comments</b>	<b>Frequencies</b>
Facilitating revision and writing development	20
Preference for specific feedback and feedback on content and organization	8
Concerns for feedback quantity and quality	5
Total	33

As can be seen from Table 6.8, the majority of the participants asserted they learned from providing feedback for their peers.

**Excerpt 6.35**

There were some good arguments in my peers' compositions. Some of their writing strategies were also worth learning. (Participant #4)

**Excerpt 6.36**

I was impressed by the advanced words and sentence patterns in my peers' compositions. (Participant #11)

**Excerpt 6.37**

It was not an easy job to offer feedback on logical relations between ideas. Sometimes, it took me half an hour to work out one solution. My effort, however, was not futile. When I encountered such problems in my writing and redrafting, I did not find them difficult at all. (Participant #41)

From the above excerpts it is apparent that students' learning from providing feedback for their peers varied. Both Participant 4 and Participant 11 learned from the strengths in their peers' compositions, specifically argumentation, vocabulary, and sentence patterns; Participant 41 learned from detecting a problem in his peer's composition and providing a solution.

Some participants also mentioned that giving feedback fostered their audience awareness which assisted in their composing and revising process.

**Excerpt 6.38**

In the role of a reviewer, I knew what other people were looking for when they were reading my compositions. Including a topic sentence within each



paragraph makes it easier for readers to get my point; presenting my arguments in separate paragraphs makes my composition more readable; a lengthy introductory paragraph and a short concluding paragraph make the text structure imbalanced. Without the experience of reading my peers' compositions, I could not realize the importance of these aspects. (Participant #25)

### **Excerpt 6.39**

I think reviewing my peers' compositions helped me recognize some problems in my writing. For instance, I found some sentences difficult to understand, probably because my peers intentionally used some complex structures. Such problems were also present in my writing. Afterwards, when I wrote some long and complex sentences, I reminded myself to check whether they conveyed my meaning clearly. (Sun, Interview)

In the above excerpts, Participant 25 identified three insights that she drew from reading his peers' compositions while Sun explained how she utilized such insights in her own writing. The treatment group participants' role as a reader fostered their awareness of readers' needs, which influenced their own writing, and led to improvements in equivalent areas.

As well as learning new ideas and specific skills and developing audience awareness, three participants claimed that giving feedback motivates them to improve their own writing.

**Excerpt 6.40**

I benefit most from the reviewing process. My original goal of participating in this research project was to have the opportunity to read compositions written by my classmates. After two or three writing tasks, I realized the intervention afforded us much more than what I had expected. While reviewing my peers' compositions, I realized the gaps between us as well as action I should take to close the gap. (Participant #13)

**Excerpt 6.41**

Reading compositions written by peers brought changes in my writing. At the beginning of the semester, I used simple structures that I was familiar with. Some of my peers, however, used complex sentence structures in their compositions. Seeing the difference, I decided to incorporate some complex structures in my writing. It was difficult at the beginning, since I needed to consult grammar books and the teacher's slides. During that period of time, the number of errors increased in my writing. Now, it is much better. When I am writing, I naturally recall some complex structures that I could use. (Participant #29)

Specifically, Participant 13 noted a change in her mentality from seeing how others performed in writing to taking action to catch up with others; Participant 29 kept on practicing using complex structures in writing, despite her awareness that using complex sentences involved effort and threatened her accuracy in writing. For the two participants, since giving feedback triggers the participants' intentional and proactive efforts in learning, improvements could be expected in due course.

As regarding receiving feedback, over half of the treatment group participants asserted that they received constructive feedback from their peers which was helpful for their revision and development in writing performance (see Table 6.9).

**Excerpt 6.42**

My classmates could always identify inappropriate words and some Chinglish expressions in my compositions. It is easier for a person to point out problems in others' compositions than in his/her own compositions. (Participant #10)

**Excerpt 6.43**

Some comments from my peers were engraved in my mind. For example, the topic sentence within each body paragraph should be concise, and descriptions about a picture should not be verbose. (Participant #34)

**Excerpt 6.44**

Some minor problems recurred in my compositions. Since my peers repeatedly pointed them out, I started to pay attention to them when writing on a new topic. (Participant #20)

In the above excerpts, the participants shared the opinion that peer feedback helped them notice problems they were unable to detect by themselves. Such feedback points naturally led to revisions which enhanced the quality of their second drafts. When participants, such as Participant 20, applied some feedback points to subsequent tasks, they avoided similar problems in their writing. Similarly, as reported in 6.2.1 and 6.2.2, students attributed their improvements in content, organization, and language use of writing to feedback received from their peers.

In addition, some participants expressed their preference for feedback on content and organization and specific feedback.

**Excerpt 6.45**

I benefitted most from feedback on thesis statement and paragraph structure.

Once I knew I had problems with these issues, I paid attention to them in subsequent writing tasks. (Participant #1)

**Excerpt 6.46**

Comments from my peers were meticulous. They marked problems on my drafts and recorded them in PF forms. Some of my peers included short but clear explanations in their comments. (Participant #15)

**Excerpt 6.47**

Most of my peers not only pointed out my problems in grammar, sentence structure, organization and content, but also included revision suggestions. Their comments were helpful. (Participant #26)

In Excerpt 6.45, Participant 1 indicated his preference for feedback on global features of writing as well as the reasons. In Excerpts 6.46 and 6.47, the participants expressed their appreciation of specific feedback in which their peers not only pointed out their problems, but also included explanations and possibly solutions to their problems. Their comments, which did not relate directly to the questionnaire item, suggest that participants were concerned about the impact of peer feedback. Because specific feedback was easily incorporated in revision, and feedback on global features showed

greater promise of improving text quality, students expressed a preference for feedback on content and organization as well as specific feedback.

Some comments, however, suggest that the treatment group participants were concerned about peer feedback quality. They identified some peer feedback as unspecific or inaccurate.

**Excerpt 6.48**

My peers pointed out some problems without explaining why. Such feedback was not helpful at all. (Participant #7)

**Excerpt 6.49**

The majority of comments from my peers were reliable, but I detected some inaccurate ones. I was worried that there were some that I failed to detect. (Participant #35)

**6.3.3 Paired discussions**

As described in 3.5.3, with each writing task, there was a 30-minute discussion session in which the treatment group participants discussed with their peers the written feedback points that they received and provided. Therefore, one item in the questionnaire concerned participants' perceived helpfulness of paired discussions. Table 6.10 presents the topics identified from the treatment group participants' comments.

Table 6.10 Students' perceptions of paired discussions

<b>Comments</b>	<b>Frequencies</b>
Supplementing written feedback	15
Creating collaborative learning opportunities	7
Obtaining teacher's help	2
Repetition of written feedback	5
Time constraint for paired discussions	3
Total	34

As can be seen from Table 6.10, most of the participants reported that the discussions supplemented written feedback.

**Excerpt 6.50**

Oral discussions were prompt and specific. For comments I felt confused about, I explained my intention and asked my peers to provide rationales for their comments. Because of my peers' explanations, I developed a better understanding about these comments. (Participant #10)

**Excerpt 6.51**

In discussions, I explained my comments in detail to my peers, and at the same time asked them to explain their intention. In some situations, I made adjustment to the written comments. My comments, are therefore, more feasible. (Participant #13)

In Excerpt 6.50, the participant, as a writer uncertain about the meaning of some feedback points, explained his own intention, and asked his reviewer to explain some feedback points, which resulted in a better understanding. In Excerpt 6.51, the participant, as a reviewer, elaborated and adjusted some feedback points to make them more appropriate for the writer. The complementary role played by paired discussions, as shown in the two excerpts, was crucial for revision. Some feedback points, if not focused on in discussions, would be likely to be ignored in revision, as shown in the following excerpts,

**Excerpt 6.52**

It is necessary to discuss the written feedback. If you received some written feedback points which you did not agree with, you might ignore them when revising. However, when you discussed them with your peers, you may change your mind. I remembered one instance in which my peer corrected my erroneous use with “a lot”. I had been using “a lot good” for so many years that I never thought it was wrong. In the discussion, my peer presented me with what he found from the dictionary. His feedback, especially his explanations in the discussion, rectified my misconceptions. Without the discussion, I will probably still use “a lot good” in my writing. (Yang, Interview)

**Excerpt 6.53**

When I was making revisions, I realized the importance of discussions. For some written feedback points which my peers and I unintentionally missed in our discussions, it took me a very long time to figure out a solution. In some

instances, I even skipped them because I could not come up with a satisfactory solution. (Wu, Interview)

Some treatment group participants also mentioned the collaborative learning opportunities in paired discussions.

**Excerpt 6.54**

When my peers and I had disagreements, the discussion was extremely helpful. We re-examined the problem together, possibly consulted some reference books or sought help from the teacher, and finally came up with a solution. It does not matter whose opinion is right. Both of us benefit from the discussion. (Participant #4)

**Excerpt 6.55**

I remembered a very fruitful discussion about a sentence. When I was writing the sentence in my first draft, I felt it a little weird but that was the only one I could come up with at that moment. My peer underlined that sentence, but she neither explained the problem nor suggested the changes. In the discussion, she told me that she did not know my intended meaning and therefore had no idea about how to revise it. When I explained to her my intension and provided some alternative words, she reminded me to think about their collocations as well as their compatibility with the rest of the sentence. We compared different ways of organizing the sentence and finally reached an agreement. In the discussion, we collaborated with each other and solved the problem through our joint efforts. (Shen, Interview)



Both excerpts highlight the co-construction of knowledge through collaboration between the reviewer and the writer. In Excerpt 6.54, the discussion enabled the two parties to resolve their confusion about a problem. In Excerpt 6.55, Shen and his peer, through clarification, negotiation and collaboration, worked out a solution that neither could achieve alone. As shown in Excerpt 6.54, the significance of paired discussions is in creating learning opportunities for both parties rather than getting a correct feedback point only for the writer, which is clearly stated by the participant.

Moreover, two participants perceived the discussions to be useful in that they could get help from their teacher. Both Participant 16 and 29 mentioned that they sought help from their teacher when they disagreed about some feedback points. Such comments indicate students' need for teacher support in peer feedback activities.

However, the treatment group participants identified two problems with paired discussions. Five participants noted that some of their peers simply repeated their written feedback, rendering the discussions unhelpful and redundant. Three students mentioned that the time constraint prevented them from having an in-depth discussion about some feedback. These comments, taken together, suggest the participants' concern for the quality of discussions.

#### **6.3.4 PF forms**

In this study, PF forms with varying foci were used for the six writing tasks. Table 6.11 presents the four topics identified from the treatment group participants' comments to the questionnaire item on the use of PF forms in peer feedback activities.

Table 6.11 Students' perceptions of using peer feedback forms

<b>Comments</b>	<b>Frequencies</b>
Facilitating provision of feedback	16
Providing guidance on writing	4
Problems in using PF form	13
Total	33

As seen from Table 6.11, almost half the participants asserted that PF forms facilitated provision of feedback for their peers. Specifically, PF forms enabled them to know what and how to comment on their peers' compositions, as well as helped them organize the written comments, as evident in the following excerpts.

**Excerpt 6.56**

Each form was helpful. From it, I knew what to focus on for each writing task and how to provide feedback. Within each form, the teacher highlighted a feedback focus by providing an introduction to it and leaving more space for corresponding feedback. For instance, when the feedback focus was on text structure, the teacher explained thesis statement and its position within a composition, the relation between the arguments and the thesis statement, and language features for thesis statement and topic sentences in body paragraphs. Such information strengthened what I had learned from the general training session and gave me guidance on evaluating this aspect for my peers. (Participant #4)

**Excerpt 6.57**

PF forms guided my evaluation of my peers' compositions. They had wide coverage of aspects to consider, including content, organization, language and vocabulary. Using PF forms prompted me to pay attention to both local issues (i.e. grammar and vocabulary) and global issues (i.e. content and organization); otherwise, I might only pick some grammatical mistakes. (Participant #21)

**Excerpt 6.58**

PF forms were useful, because they helped me organize the problems I identified from my peers' compositions into categories. With PF forms, I would know which aspects I missed when providing feedback. (Participant #18)

In addition, four participants stated that PF forms guided their writing because through constant use of PF forms, they became familiar with, and consequently give consideration to these criteria, when writing.

**Excerpt 6.59**

The criteria in PF forms are the qualities that a good piece of expository writing should have. (Participant #13)

**Excerpt 6.60**

After using PF forms for several tasks, the criteria for evaluating a composition naturally came into my mind when I was writing. The introductory paragraph should include a summative and concise thesis

statement; the arguments should not overlap; sentences and paragraphs should be coherent; sentence structures and words should be various; grammatical mistakes should be avoided. I was very familiar with these criteria. Although I could not perform satisfactorily in all these aspects, they still had some influence on my writing. (Liu, interview)

Thirteen participants, however, identified problems with using PF forms. They expressed their doubt about the need to fill PF forms and showed their disapproval of some content in and requirements of the PF forms, as seen in the following excerpts.

**Excerpt 6.61**

The PF forms guided my evaluation of peers' compositions, but I preferred to write down my comments on the drafts in that such comments were helpful for the discussion and revision. (Participant #14)

**Excerpt 6.62**

Filling in PF forms doubled my workload and it was less effective than marking and writing my comments on the first drafts. (Participant #28)

**Excerpt 6.63**

In paired discussions, we referred to the comments that we marked on the drafts. If we followed the PF form, both parties had to locate the problems in the drafts. It wasted time. When writing my second draft, moving between the PF form and my first draft was not convenient at all. (Shen, Final interview)

#### **Excerpt 6.64**

Some criteria in PF forms were rigid. They did not fit my peers' compositions. Also, for some writing tasks, I could not generate as many comments as I was required to. Sometimes I was hypercritical about my peers' compositions because every section of the form needed to be completed. (Participant #33)

#### **Excerpt 6.65**

Some sections are unnecessary, such as "strengths of my peer's composition". I filled in this section, but I neither discussed this section with my peers in oral discussions nor considered it when revising my own composition. (Participant #40)

In Excerpts 6.61 and 6.62, the participants indicated their preference for writing their comments on the first draft because completing PF forms increased their workload with no beneficial effect. Shen further elaborated the inconvenience caused by the use of PF forms in discussions and revisions in Excerpt 6.63, and in Excerpts 6.64 and 6.65, the participants expressed their dissatisfaction with specific elements. Although none of the problems pertained to the facilitating role that PF forms played in the provision of feedback, they were practical issues to consider in the use of PF forms.

### **6.4 Chapter summary**

This chapter has reported the treatment group participants' perceptions of peer feedback, based on the 35 treatment group participants' answers to the Perceptions of Peer Feedback Questionnaire and the six case study participants' responses in the semi-structured interviews.

The treatment group participants reported their improvements in content, organization and language use as a result of peer feedback activities. They described the specific areas of content and organization in which improvements had occurred and explained how they made the improvements. The majority of the participants also assumed their use of language improved: Their accuracy increased; their grammatical knowledge was enhanced; their vocabulary expanded. Some participants identified their lack of improvement in language use and attributed it to the recurring nature of language problems and their avoidance of complex structures in writing.

Similarly, most of the treatment group participants perceived the six components of the intervention to be helpful. The excerpts selected from questionnaires and interviews illustrated how each component functioned within the intervention. The participants confirmed that the general training session prepared them for participating in the intervention and fostered self-regulation behaviors in writing, and that the focused training sessions provided them with ongoing and specific support and facilitated revision. They reported that giving feedback helped them learn new ideas and specific skills, train reader awareness, and provide motivational impetus. Participants acknowledged that receiving feedback facilitated revision and writing development. They expressed their preference for feedback on global aspects of writing and specific feedback. Paired discussions were described as useful in that they supplemented the written feedback that the participants received, created collaborative learning opportunities, and involved their teachers in peer feedback activities. Participants agreed that PF forms served as a guide when providing feedback for their peers and when writing on a new topic. While participants agreed that each component was useful, they identified some problems. They were concerned about the quality of their received

feedback, the time constraints and quality of the paired discussions, and the use of PF forms; these should be considered when reviewing the effectiveness of the intervention.

## **Chapter 7 Discussion**

### **7.1 Chapter overview**

This chapter discusses the results reported in the preceding three chapters. It starts with discussing the effects of peer feedback on text revisions and the effects of peer feedback on writing performance, integrating the quantitative and qualitative results in relation to the findings and argument of previous studies. The next two sections discuss how the design of the intervention and the participants' efforts in utilizing peer feedback in revision contribute to the efficacy of peer feedback. Finally, an adapted Hayes' (1996) revision model is presented, based on the findings from this study and previous studies on feedback.

### **7.2 Effects of peer feedback on text revisions**

The quantitative data collected from the Writing Task 6 in the quasi-experimental study and the qualitative think-aloud data from the case study participants address the effects of peer feedback on text revisions. The results suggest that the peer feedback intervention implemented in this study was more effective than collective feedback in helping students improve their drafts, particularly in global aspects of the written texts.

#### **7.2.1 Overall quality**

Both the treatment group and comparison group made significant improvements in text quality from their first drafts to their second drafts, with the treatment group significantly exceeding the comparison group in the overall score for their second drafts.

The enhanced overall quality of the treatment group participants' second drafts is consistent with the findings in existing L2 studies that involved only one group of



participants receiving peer feedback and used the overall score, the incorporation rate of feedback, or frequency of self-revisions, to measure the effects of peer feedback on text revisions. (Hu & Lam, 2010; Kamimura 2006; Mendonça & Johnson, 1994; Min, 2006; Villamil & de Guerrero, 1996). In the present study, this result can be also inferred from the analysis of feedback points that the case study participants received and attended to when writing their second drafts. Given that the case study participants received feedback on content, organization, and language use (see Table 5.2) and responded to 82% of the received feedback points when writing their second drafts of Writing Task 6 (see Table 5.3), an improvement in the second drafts would be expected. The high take-up rate of peer feedback, however, needs to be interpreted with caution. The percentages reported in previous studies (Hu & Lam, 2010; Mendonça & Johnson, 1994; Villamil & de Guerrero, 1996) were obtained from text analysis of changes and received peer feedback points, so only changes present in revised drafts were taken as evidence for incorporation. In other words, some instances in which the participants made the decision not to use the received feedback points were missing from the data. In this study, all feedback points that the case study participants attended to, as observed in the think-aloud data, were considered as evidence of uptake, regardless of whether they were incorporated in their second drafts.

While some L1 studies compared peer feedback with revision training or self-revision, the comparative effects of peer feedback and teacher feedback were mainly from L2 studies, with divergent findings (Connor & Asenavage, 1994; Paulus, 1999; Tsui & Ng, 2000; Ruegg, 2015b; M. Yang et al., 2006; Zhao, 2010). The finding from this study that peer feedback had an advantage over collective feedback on text revisions was supported by studies that reported the success rate of revisions, percentage of meaning changes and understanding rate of feedback (e.g., Paulus, 1999; Ruegg, 2015b; Zhao,

2010). Studies using writing scores, the incorporation rate of feedback, and the percentage of revisions triggered by feedback concluded that peer feedback was less effective than teacher feedback for text revisions (e.g., Connor & Asenavage, 1994; M. Tsui & Ng, 2000; Yang et al., 2006). In my study, the fact that the peer feedback mode was both written and oral might be a factor leading to the treatment group participants' better performance in text revisions. When only oral peer feedback was implemented, as in Connor and Asenavage (1994) and Mendonça and Johnson (1994), the participants, without any preparation, may have received insufficient feedback for revision. The participants in Ruegg (2015b), because of the lack of oral peer feedback, may have encountered some difficulty in understanding the received feedback points and in using them in revision. In addition, the focused peer feedback training session within each writing task of this study may have a tangible effect on the treatment group participants' better performance in text revisions. The L2 studies examining the effects of peer feedback on text revisions as I have reviewed, with the exception of Kamimura (2006), all collected drafts from writing tasks without peer feedback training.

### **7.2.2 Content and organization qualities**

Both groups made significant gains in content and organization scores between drafts, but the treatment group made more substantial gains in each score, as can be seen from the larger effect sizes for the treatment group than for the comparison group ( $d = .90$  vs.  $d = .34$  for content score, and  $d = .88$  vs.  $d = .36$  for organization score). The treatment group participants also outperformed those in the comparison group in the content and organization qualities of their second drafts, indicating the peer feedback intervention was more helpful than collective feedback for text revisions in content and organization.

The finding that the treatment group made significant improvements in content and organization qualities between drafts adds to the existing evidence for the positive effects of peer feedback on text revisions in content and in organization. With only one group of participants receiving peer feedback, Hu and Lam (2010) similarly reported that peer feedback helped students make gains in the content and organization scores in their second drafts and Min (2006) found that texture revisions (concerning coherence) and explicature (concerning explanation) ranked first and second in all revisions that their participants made. Olson (1990) found that of her four groups (receiving revision instruction and working with peer partners, working with peer partners only, receiving revision instruction only, and control group), only the peer feedback group made a significant improvement in rhetorical quality scores between drafts.

Findings from the case study may help explain the improvements in content and organization qualities that the treatment group participants made between drafts. As presented in 5.3.2.2, in response to feedback on content and on organization, the six case study participants mainly adopted the elaborated approach, which they implemented various cognitive operations and generated changes in multiple steps. Storch and Wigglesworth (2010) argued that extensive engagement with corrective feedback led to students' correct use of certain language forms. Similarly, it could be claimed that the elaborated approach in processing feedback on content and on organization would have the same positive impact on the qualities of content and organization of their second drafts.

The finding that the treatment group outperformed the comparison group in content and organization qualities of their second drafts was not consistent with the findings in Cho and Schunn (2007), who found that the participants in the single peer condition and in

the single expert condition had similar performance in content and organization qualities for their revised drafts. They explained that feedback points from the expert might have been unhelpful in that they were based on the expert's knowledge, beyond the level of the participants. Their explanation, however, seemed inappropriate for the finding in my study. The expert in Cho and Schunn (2007) had no experience in language teaching while the teacher in this study had many years of experience in EFL teaching. The peer feedback being specific in contrast with collective feedback being generic might be a possible reason for the treatment group participants' better performance in content and organization of their second drafts. In my study, Min's (2005) four steps in providing feedback was introduced and practiced in the general training session and the activities of distinguishing between helpful and problematic comments were frequently included in focused training sessions. The collective feedback for the comparison group differed from the conventional teacher feedback in that 1/3 to 1/2 of the participants received written feedback while the rest only received oral feedback reminding them of some common problems. Because the problems in content and in organization were difficult to detect and solve, the participants who only received oral feedback may have failed to see the relevance of some feedback points to their writing and therefore made very few revisions. In addition, the focus of PF form and training session in Writing task 6 being on cohesion, a topic in organization, may explain the treatment group participants' better performance in organization for their second drafts. The requirement of two feedback points on cohesion in the PF form ensured that each participant received some feedback points on cohesion while the instruction in cohesion in the PF form and the teacher's explanations in the training session facilitated the treatment group participants' revisions in cohesion.

### 7.2.3 Language use

Peer feedback activities impacted the four measures of language use in revision differently. Students in the treatment group improved significantly in accuracy and in syntactic complexity in their second drafts, but not in the other two measures, lexical complexity and fluency. Students in the comparison group, however, improved significantly only in accuracy between drafts. There was no significant difference between the two groups in any of the four measures for their second drafts.

Findings from the case study provide evidence of improvement that the treatment group participants made in accuracy between drafts. As presented in 5.3.2.1 and 5.3.2.2, the case study participants used the direct approach to deal with feedback points on surface problems (e.g., spelling, tense, and number), and employed the elaborated approach when responding to feedback points on lexical and sentence syntactic problems. Based on analysis of the eight case study participants' paired discussions about the received feedback points and their revised drafts, Storch and Wigglesworth (2010) argued that limited engagement with feedback on surface errors, reading the received feedback and indicating acceptance, is sufficient for revisions and that extensive engagement with feedback on demanding issues resulted in successful revisions. For the same reason, the treatment group participants in this study produced their second drafts with improved accuracy. In addition, the high frequency with which the case study participants employed the cognitive operation of monitoring (see Table 5.4), particularly when responding to feedback on content and organization (see 5.3.2.2), could also predict their improvement in accuracy. By checking the accuracy and appropriateness of the changes, the case study participants avoided making similar errors in their second drafts.

The result that the treatment group participants made a significant improvement in accuracy partially corroborates previous studies. Using a single score to evaluate multiple aspects of language use, Hu and Lam (2010) and Olson (1990) found the participants receiving peer feedback made a significant improvement in language use in their revised drafts. Diab (2010), the only study with a specific focus on accuracy, reported that students in the peer feedback condition significantly reduced in rule-based errors (subject/verb agreement and pronoun agreement errors), but not in non-rule-based errors (word choice and sentence structure errors), in their second drafts. The finding from this study, when examined together with that from Diab's (2010) study, shows that peer feedback activities helped students improve in overall accuracy and in accuracy specific error types.

The finding that peer feedback and collective feedback were equally effective in reducing errors in drafts has not been reported in the previous literature. The finding has two implications. Firstly, EFL students did not lack the ability to generate reliable feedback on language errors and make revisions; and working with peers did not disadvantage them in reducing errors in their drafts. In this study, many factors (e.g. teacher support and consulting reference materials) may have compensated for the participants' lack of proficiency in L2. Secondly, EFL students have a strong tendency to focus on language problems in revision, whether they were working with peers or receiving feedback from their teachers.

The finding that participants in the treatment group produced significantly longer T-units, evidence of greater syntactic complexity, in their second drafts, was new and encouraging. The co-occurrence of significant improvements in content and in organization, and significant increase in the mean length of T-units in this study, imply

that adjusting information in response to feedback on content and organization may result in longer ideas units. Although this study did not examine changes made by the treatment group participants in their second drafts, previous studies have shown that students receiving peer feedback made many revisions by clarifying and elaborating what they had written in their first drafts, and that such revisions occurred most at sentence level (Lam, 2013; Min,2006). The treatment group participants' improvement in syntactic complexity may also be because of their voluntary and deliberate use of complex structures in their second drafts. Since the data were collected from the last writing task in the intervention, the participants might have acquired some sentence and grammatical structures after constantly reading their peers' compositions and receiving instruction through PF forms and focused training. With no pressure of time or grades, they may have experimented using new structures leading to longer and more complex sentences in their second drafts.

The results that neither group improved in lexical complexity between drafts and both groups performed similarly in their second drafts indicate that neither treatment was effective in helping students improve in lexical complexity. A cursory reading of the written documents for the six writing tasks collected from the treatment and comparison group participants reveals that there were very few peer or teacher feedback points targeting on word choice. For each writing task, both the teacher and students pointed out some grammatical errors and word collocation problems. The narrow focus may have prevented participants in either group from making an improvement in lexical complexity between drafts. Findings from the case study may help explain the lack of improvement in lexical complexity for the treatment group. As seen from the analysis of the received feedback points and think-aloud protocols for Writing Task 6, Liu was the only case study participant who paid attention to lexical complexity when writing

her second draft. She replaced four words in her first draft with more sophisticated ones: Three were prompted by her peer (replacing “in my opinion” with “as far as I am concerned”, “great” with “significant”, and “important” with “vital”), and one was initiated by herself (replacing “wonderful” with “appealing”). Such changes, rare among the case study participants and possibly among the treatment group participants, however, might be too sporadic to bring about a significant change to lexical complexity.

Similarly, neither group changed significantly in fluency, as measured by text length, between drafts, and there was no significant difference between the two groups in fluency for their second drafts. The finding that the treatment group participants did not demonstrate an improvement in fluency between drafts was consistent with Kamimura (2006), which claimed peer feedback did not help high-proficiency Japanese EFL students’ text length increase in their revised drafts. Kamimura (2006) explained that the students’ high L2 proficiency level may have prevented them from making any further improvement in text length. The participants in the present study, in the treatment group or the comparison group, had high L2 proficiency, as seen from this performance in the university English placement test (see 3.6.1). In the present study, the specification for the minimum word count in writing task directions was another possible reason: EFL students were frequently reminded that writing excessively would affect their writing scores in exams. As the mean text length of the first draft for each group already exceeded the minimum word count by around 50 words (see Table 4.5), it was likely that participants in both groups deliberately controlled the length of their second drafts.



### **7.3 Effects of peer feedback on writing performance**

For the effects of peer feedback on writing performance, quantitative data gathered from the pre-test, post-test, and delayed post-test in the quasi-experimental study and the perception data from questionnaires and interviews are discussed together. In general, the peer feedback intervention was more effective than collective feedback in helping students improve in global aspects of writing and accuracy, and only improvements made by the treatment group was retained twelve weeks after the intervention.

#### **7.3.1 Overall quality**

The results from the three writing tests showed that both the treatment group and the comparison group made gains in overall scores in the post-test and in the delayed post-test. The treatment group, however, significantly outperformed the comparison group in overall scores in the post-test and in the delayed post-test. These results confirmed that peer feedback was more effective than collective feedback in helping students improve in overall writing quality.

The finding about the positive effects of peer feedback on the improvement in writing quality is consistent with some previous studies (e.g. Gai & Zhou, 2013; Karegianes, Pascarella, & Pflaum, 1980; Zhang & McEneaney, 2020). Some other studies reported that peer feedback group participants made similar improvements in writing quality as those in the teacher feedback group (Gielen et al. 2020; Ruegg, 2015a; Y. Zhou, 2013), and L. Yang et al. (2013) reported that traditional peer feedback (to be distinguished from online peer feedback) was less effective than teacher feedback in helping students improve the writing quality. For the discrepant findings, one possible reason may be the format of training. The present study, as Karegianes et al. (1980) and X. Zhang and McEneaney (2020), provided the participants with well-structured peer feedback

training throughout the intervention period. Conducted with L1 learners at the secondary level, Gielen et al. (2010) acknowledged that they only provided “minimal” training through a modeling activity (p. 151). Similarly, Ruegg (2015a) provided her L2 students at the tertiary level with training only prior to the intervention. Another reason may concern the length of the intervention, about which Prater and Bermudez (1993) explained that “one month is not a sufficient length of time to produce significant differences in overall quality of writing” (p. 108). The present study, with a 13-week intervention and six writing tasks with feedback activities, featured a relatively long intervention period with frequent feedback iterations. In this study, the better effects of peer feedback on the improvement in writing quality may also be attributed to PF forms, specifically the instruction sections. With the purpose of facilitating evaluation of their peers’ compositions, such PF forms may have helped students strengthen their writing knowledge.

### **7.3.2 Content and organization**

The results from the quasi-experimental study indicate that the peer feedback intervention had positive effects on improving content and organization qualities in writing. Only participants in the treatment group significantly progressed in content and organization scores over time. The treatment group had higher content scores than the comparison group in the post-test and delayed post-test, but neither difference was statistically significant. The treatment group significantly outperformed the comparison group participants in organization scores in the post-test and delayed post-test, suggesting the peer feedback intervention was more effective than collective feedback in helping students improve organization qualities in writing.

The perception data from questionnaires and interviews corroborate the quantitative findings in showing that the peer feedback intervention was helpful in improving students' performance in the content and organization of writing. Among the treatment group participants, a high percentage opted for "agree" with the statements about their improvements in content (94.3%) and organization (94.3%), clearly and confidently identifying areas in content and organization that they felt had improved. Since these areas resemble common criteria for evaluating a composition, their responses suggest that they had developed an understanding of expectations for content and organization in their compositions. Consequently, they may have paid attention to these expectations when writing on the post-test topic, resulting in texts with better quality in content and in organization.

Also evident in the perception data is that improvement in text structure was mentioned most frequently by the treatment group participants and the case study participants. A comparison of compositions written by the treatment group participants for the pre-test and post-test provides further evidence for this finding. In the pre-test, the treatment group participants commonly used the three-paragraph structure, with a long body paragraph including several arguments and a short concluding paragraph with only one sentence. Their post-test compositions, however, were more balanced in text structure, with two or three body paragraphs and a concluding paragraph including two or three sentences. Such findings were confirmed by the large effect size indicating the magnitude of differences between pre-test and post-test organization scores ( $d = 1.25$ ) for the treatment group.

The result that peer feedback benefited content and organization qualities in writing has been reported in studies that compared the effects of giving feedback and receiving

feedback. The three studies (Crinon, 2012; Crinon & Marin, 2010; Lundstrom & Baker, 2009) reported that both feedback giving group and feedback receiving group improved in the overall, content and organization qualities of their writing. Feedback giving group, however, improved more than feedback receiving group in each of the three aspects.

The finding that there was no significant difference between the two groups in the content quality of their post-test compositions is consistent with findings reported in Ruegg (2015a) and Y. Zhou (2013), who also included a sub-score for assessing the content. The finding from the present study, however, was different from that of L. Yang et al. (2013), who found students receiving teacher feedback outperformed those receiving peer feedback in the content quality of writing. The discrepancy may be due to the difference in treatments. The teacher feedback group in L. Yang et al. (2013) received individual written feedback on drafts and collective feedback addressing common problems in class, in contrast with what was carried out in my study, where only collective feedback was provided to the comparison group. The combined use of individual written feedback and collective feedback may have led to the teacher feedback group participants' better performance in content quality of writing.

The finding that the treatment group had a higher organization score in the post-test than the comparison group was not consistent with what was reported in Ruegg (2015a), who reported that students in the peer feedback group made similar gains in the organization score as those in the teacher feedback group in the post-test. The current finding was in contrast with Y. Zhou (2013) and L. Yang et al. (2013), both of whom found the teacher feedback group made greater gains as shown in their organization scores in the post-test than the traditional peer feedback group. While the treatment group participants in my study received written comments and had time to discuss them, all the participants in

Ruegg's (2015a) study received written feedback through feedback forms; and the traditional peer feedback group participants in Y. Zhou's(2013) study received oral feedback through group discussions. The combined use of written and oral feedback in my study may have helped the treatment group participants receive sufficient feedback points and enhance their understanding about them, particularly those on the content and the organization. Because feedback on organization points (e.g., text/paragraph structure and cohesive devices) could be applied to other writing tasks, unlike the topic-dependent content feedback points, students may have little difficulty using them when working on a new task. For the better effects of teacher feedback on the organization quality in writing reported in L. Yang et al. (2013), the combined use of individual written feedback and collective feedback, as mentioned above, might be a possible reason.

### 7.3.3 Language use

The peer feedback intervention impacted the four measures of language use differently, as evident in the findings from the quasi-experimental study. Both the treatment group and the comparison group improved significantly in accuracy in the post-test, but only the treatment group retained the improvement in the delayed post-test. The treatment group significantly outperformed the comparison group in accuracy only in the delayed post-test. In syntactic and lexical complexity, neither group had any significant change, suggesting that neither the peer feedback intervention nor collective feedback had effects on students' performance in lexical or syntactic complexity of writing. With fluency, only the treatment group made significant gains in text length from pre-test to post-test, but the two groups performed similarly in the post-test and delayed post-test. Such results suggest that while the peer feedback intervention had a positive effect on students' fluency in writing, its effect was not retained in the delayed post-test.

The result of positive effects of peer feedback on accuracy is corroborated by the perception data obtained from questionnaires and interviews. Although clearly stated in the questionnaire that language use included grammar, vocabulary, and sentence structure, more than half of the treatment group participants cited their decrease in errors as evidence for their perceived improvement in language use. The high frequency with which improved performance in accuracy was mentioned indicates that the treatment group participants, as commonly observed among EFL writers, accorded a high priority to accuracy in their writing.

The finding that the treatment group participants had better performance in accuracy in the delayed post-test than the comparison group is not incongruent with Ruegg (2015a) or L. Yang et al. (2013) who reported peer feedback was less effective than teacher feedback in improving students' performance in accuracy, as measured by a grammar score. Regarding the smaller improvement for the peer feedback group participants, Ruegg (2015a) explained that some feedback on surface-level errors given by peers may have been inaccurate and therefore ineffective. In the present study, the quality of peer feedback on language problems was not a significant issue. Not only did the teacher present common grammatical problems selected from the participants' compositions during a focused training session, she also constantly reminded the participants to consult reference resources when generating and utilizing feedback on language issues. The treatment group participants' use of reference books when providing feedback and writing their second drafts may have created valuable opportunities for language learning which the comparison group did not have.

With syntactic and lexical complexity, the findings were somewhat surprising because both groups remained unaffected by the intervention. Existing peer feedback studies

typically used a grammar score to measure accuracy and sentence structure varieties at the same time and a vocabulary score to measure lexical range, appropriateness and sophistication (e.g., Hu & Lam, 2010; Ruegg, 2015a; Y. Zhou, 2013; L. Yang et al., 2013); no peer feedback studies, as far as I know, have employed the syntactic complexity index of MLT and lexical complexity of MSTTR\_50. The current finding of all participants' lack of significant change in syntactic complexity may suggest that under the stressful examination conditions, students consider content, organization and/or accuracy, but pay no attention to syntactic complexity.

Previous studies reported divergent findings on the effects of peer feedback on vocabulary in writing: Ruegg (2015a) reported that students receiving peer feedback made similar gains as those receiving teacher feedback in vocabulary score; Lundstrom and Baker (2009) found that neither the feedback giver group nor the feedback receiver group made significant gains in vocabulary score after a semester-long intervention; Y. Zhou (2013) found the traditional peer feedback group and online peer feedback group made greater gains in vocabulary score than the teacher feedback group; L. Yang et al. (2013) showed that the peer feedback group participants made smaller gains in vocabulary score than the other two groups, receiving teacher feedback alone or combined teacher and peer feedback. A possible reason for the discrepant findings about vocabulary might be that mean segmental type token ratio only measures diversity of words whereas scoring rubrics employed in the above-mentioned studies included other criteria, such as lexical appropriateness and sophistication. Another possible explanation might be duration of the intervention: The intervention periods in the present study, in Lundstrom and Baker (2009), and in Yang et al. (2013), were one semester, whereas the intervention in Ruegg (2015a) and Y. Zhou lasted for one academic year. It is possible that one semester is inadequate for students to attain substantial developments in

vocabulary knowledge. Also, the perception data reveal another reason for the lack of improvement in lexical complexity for the treatment group: the difficulty in using the words learned from peer feedback activities. Because the newly acquired words are in the form of explicit knowledge, which students cannot retrieve when writing on a new topic, an improvement in lexical complexity is unlikely. However, if teachers intervene and design some activities to help students rehearse the newly-acquired words to facilitate the transfer into implicit knowledge, students may improve in lexical complexity in due course.

With fluency, the finding that only the treatment group participants increased significantly in text length in the post-test is consistent with some previous studies (Crinon, 2012; Crinon & Marin, 2010; Prater & Bermudez, 1993; Shang, 2019). In the present study, the change in text structure between pre-test and post-test for the treatment group participants (in 7.2.2) helps explain the significant increase in text length in the post-test. When a body paragraph is developed into two or three body paragraphs, each with some added supporting details, compositions naturally increase in text length. The possible reasons for the similar performance in fluency of the two groups in the post-test and in the delayed post-test, are the time constraints of tests and minimum word requirement in the writing directions. With the mindset that writing more than necessary may increase the chance of making errors and negatively influence their writing quality, EFL learners tend to adhere strictly to the minimum specified word requirement in examination conditions.

Apart from the four sub-dimensions of language use, the perception data revealed two areas that peer feedback may influence positively: L2 knowledge and linguistic awareness. Some treatment group participants reported that they strengthened



understanding about grammatical structures, expanded their vocabulary, and developed the habit of self-editing errors and considering word and sentence varieties. Similarly, Min (2005) argued that peer feedback activities aided language acquisition and writing development. Both improved L2 knowledge and greater attention to language use have the potential for improved performance in language use in writing: The former provides students with rich language resources while the latter helps students monitor their language use when composing.

#### **7.3.4 Retention of effects for peer feedback on writing performance**

The effects of peer feedback on the retention of gains, evident in the participants' performance in some measures in the delayed test, twelve weeks after the intervention, are important findings from the present study. They provide robust evidence for effectiveness of the peer feedback intervention. To facilitate the discussion of this retention of effects, the pre-test and post-test, the pre-test and delayed post-test results are all presented in Table 7.1.

As shown in Table 7.1, both the treatment group and comparison group retained their improvements in overall writing quality in the delayed post-test. The treatment group participants made significant improvements in written content, organization, accuracy, and fluency in the post-test and they retained their improvements in written content, organization and accuracy in the delayed post-test. The comparison group participants only made significant improvements in accuracy in the post-test, but they failed to retain their improvement in the delayed post.

Table 7.1 Changes between pre-test and post-test, and between pre-test and delayed post-test

	Treatment group		Comparison group	
	Pre-->Post	Pre--> Del	Pre-->Post	Pre -->Del
<b>Overall</b>	↑	↑	↑	↑
<b>Content</b>	↑	↑	--	--
<b>Organization</b>	↑	↑	--	--
<b>Accuracy</b>	↑	↑	↑	--
<b>Syntactic complexity</b>	--	--	--	--
<b>Lexical complexity</b>	--	--	--	--
<b>Fluency</b>	↑	--	--	--

*Note.* Upward arrow = significant improvement; Dash = no significant change; Pre = Pre-test; Post = Post-test; Del = Delayed post-test.

The findings of the treatment group participants' improved performance in written content, organization and accuracy in the delayed post-test suggest that the participants internalized what they gained from peer feedback activities and automatically applied such knowledge when working on a new topic. With only a post-test immediately after the intervention, we may lack the confidence to claim the participants' gains, as seen from their compositions in the post-test, were the results of their internalized knowledge. The participants may have paid attention to some problems just because their peers constantly reminded them to do so. For the comparison group participants, their regression in accuracy in the delayed post-test may suggest their over-reliance on teacher feedback. In the post-test, they may have paid attention to some language problems, as repeatedly pointed out by their teachers. After some independent writing tasks without feedback, they gave little attention to such issues when writing.

## **7.4 Effectiveness of peer feedback intervention**

The positive impact of peer feedback on text revisions and on writing performance may be an outcome of the intervention. The design of peer feedback intervention in the present study was based on social constructivism with particular attention to peer collaboration and teacher scaffolding. The four-month intervention provided EFL learners with ample opportunities to interact with multiple peers through the exchange of written and oral feedback and to receive teacher scaffolding in a sustained way.

### **7.4.1 Reciprocal peer feedback activities**

Reciprocal peer feedback activities, in which students give as well as receive feedback, was a major contributing factor to the success of the intervention. Although previous studies have reported that students giving feedback made greater improvements in writing performance than those receiving feedback, particularly in the global aspects of writing, it is agreed that reciprocal peer feedback activities are needed (Crinon, 2012; Crinon & Marin, 2010; Lundstrom & Baker, 2009; Cho & MacArthur, 2011). Reciprocal peer feedback activities can help learners “develop the thinking skills necessary to effectively evaluate a paper, as well as practice using feedback they receive from their peers” (Lundstrom & Baker, 2009, p. 39).

The comments from the treatment group participants reveal distinctive roles played by giving feedback and receiving feedback. In previous studies, the contributions of giving feedback have been widely reported, including improved writing proficiency in terms of content and organization qualities, opportunities for language learning, and reader awareness (Cho & MacArthur, 2011; Crinon, 2012; Crinon & Marin, 2010; Lundstrom & Baker, 2009; Min, 2005; Tsui & Ng, 2000). The treatment group participants in this study also recognized the value of giving feedback in learning new ideas or specific

skills and developing reader awareness. In the role of feedback giver, they obtained new ideas from reading their peers' compositions and learned new skills by "detecting and diagnosing problems", "recommending solutions", and "giving explanations" (Cho & MacArthur, 2011, p. 75). When students pointed out problems in their peers' writing, they became more sensitive to similar problems in their own writing. In addition, the comments from the treatment group participants reveal another contribution of giving feedback, providing motivational impetus, a contribution which has not been reported in previous studies. When reviewing their peers' compositions, the participants compared their own comparisons with those written by his peers and appeared to take the opportunity to emulate them. Their proactive efforts would likely lead to improved performance in writing.

The treatment group participants perceived receiving peer feedback to be an essential component in the intervention and recognized its value in facilitating revision and writing development. Previous studies have noted the value of receiving specific comments only for revision, helping writers identify problems and produce a revised draft with enhanced quality (e.g., Min, 2005; Lockhart & Ng, 1993, Mendonça & Johnson, 1994). Similarly, the treatment group participants in this study mentioned repeatedly that their peers pointed out problems they were unable to detect themselves, showing the indispensable role that receiving feedback played in revision. Also, some of them indicated that they applied some learning to future writing tasks, thus extending the value of receiving feedback beyond facilitating revision.

#### 7.4.2 Paired discussions

The paired discussions in the design also contributed to the success of peer feedback intervention. Within each writing task, the treatment group participants had 30 minutes to discuss written feedback and their first drafts with their peers during class time.

From the socio-cultural perspective, social interaction contributes to language development in that it “produces new, elaborate, advanced psychological processes that are not available to the organism working in isolation” (Vygotsky, 1989, p. 61). Empirically, some studies have argued that collaborative dialogues between peers, as a typical form of social interaction, contributes to language development in general and writing ability in particular (de Guerrero & Villamil, 2000; H. H. Li et al., 2020; Ohta, 2000). In peer dialogues, the two participants jointly work out solutions to linguistic problems and build knowledge which “may become a tool for their individual use of their second language” (Swain, 2000, p. 104). Similarly, the treatment group participants in the present study reported in the questionnaires that paired discussions provided them with opportunities to co-construct knowledge by jointly analyzing and solving problems in their first drafts.

Another advantage of paired discussions, as reported by the treatment group participants, was that it made the written comments more specific and applicable in revision. Without paired discussions, only reviewers were responsible for making written feedback specific, by clarifying the writer’s intention, explaining a problem, and/or providing a suggestion (Cho & Schunn, 2007; Min, 2005). In paired discussions, the writers’ efforts, however, seemed to contribute to feedback specificity through clarifying their intentions, asking the reviewers to elaborate their comments, as well as soliciting suggestions from the reviewers. The think-aloud data showed that some feedback points discussed by the

reviewers and the writers were incorporated easily in revision; whereas some feedback points that were not discussed, particularly those on global aspects of writing, required more mental effort from the writer or were neglected in revision.

Some treatment group participants indicated that the availability of teacher support was an advantage and acknowledged the need for teacher intervention in peer feedback practice, as will be discussed in the following paragraphs.

#### **7.4.3 Various forms of teacher support**

Another factor contributing to the success of the peer feedback intervention in this study was teacher support in various forms, which were implemented through a general training session prior to the six writing tasks, six focused training sessions, and PF forms. From a sociocultural perspective, teacher support constitutes scaffolding, which functions through an expert controlling the elements of a task to help a learner concentrate on elements within his/her range of competence and achieve a higher level of competence (Wood, Bruner, & Ross, 1976; L. J. Zhang & Cheng, 2020).

The general training in this study differed from that from previous studies in its focuses. Apart from focusing on feedback and interaction strategies as previous studies (e.g., Stanley, 1992; Min, 2006; Rahimi, 2013), the general training in this study had a focus on revision strategies. Specifically, the general introduction to peer feedback aimed at “getting students interested in a task and helping them adhere to the requirements of the task”, and the instruction and modelling of feedback, interaction and revision strategies aimed at “maintaining the learner’s pursuit of a particular objective” (van de Pol, Volman, & Beishuizen, 2010, p. 276). The treatment group participants’ responses in the questionnaires suggest that the general training prepared them cognitively and affectively for upcoming peer feedback activities. In addition, quite a number of the

treatment group participants indicated that they applied what they gained from the general training in their own writing through self-monitoring and self-evaluation, which positively impacted the quality of their compositions. Previous studies also found that peer feedback training helped students incorporate feedback points and make more revisions with enhanced quality, as well as improve in overall writing ability (Min, 2006; Rahimi, 2013).

The focused training sessions, implemented regularly during the intervention period, addressed challenges that the treatment group participants faced in actual feedback practices, particularly during revision. As the teacher focused on a specific skill in each session and selected authentic feedback points that treatment group participants generated for their peers, she displayed the scaffolding behaviours of “reduction of degrees of freedom” and “marking critical features” by regulating the peer feedback task to a level appropriate to the participants (Wood et al., 1976, p. 98). While no existing study has adopted such a sustained and focused training approach, some studies have indicated the need to provide students with ongoing support in addition to the pre-intervention training (Lee, 2015; X. Zhang & McEneaney, 2020; Q. Zhu, 2018). The treatment group participants’ comments in the questionnaires suggest that they appreciated the specific focus in each training session and acknowledged its role in supplementing peer feedback and facilitating revision. The focused training sessions helped the participants notice problems that their peers failed to identify and understand problems that their peers were unable to explain convincingly. To some extent, the focused training sessions played a role similar to teacher feedback on peer feedback, which “address[ed] students’ concerns over its [peer feedback] validity and prevent[ed] learners misleading each other with incorrect peer feedback” (Zhao, 2014, p. 157) .

The PF forms which the researcher and the teacher jointly designed to guide students' provision of feedback, also fulfilled their expected functions. To achieve the optimal effect of peer feedback, Q. Zhu (2018), based on her observation data and students' self-accounts, highlighted the need for pre-emptive teacher guidance within each writing task. Given that the tight curriculum did not allow the teacher to provide pre-emptive guidance, the present study employed PF forms which included instruction for a specific feedback strategy as an alternative, fulfilling the scaffolding function of "demonstration" (Wood et al., 1976, p. 98). It was apparent from the questionnaire responses that, although some participants indicated their disapproval of the compulsory use of PF forms and of some content in PF forms, they acknowledged that PF forms provided them with guidance in generating feedback. In addition, in contrast to the finding in W. Wang (2014) that using PF forms for six writing tasks within an academic year resulted in "instrumentalism", "leading students to write to the rubric to achieve a high score" (p.93), some treatment group participants in this study reported that consistent use of PF forms enhanced their understanding of good writing and positively influenced their own writing.

#### 7.5 Students' efforts in processing peer feedback

The previous section describes several design features in this study that may have contributed to the positive impact of peer feedback on revision and writing performance. It should be noted, however, that peer feedback does not automatically lead to improved performance, which depends on how students make use of the feedback they receive. When students indiscriminately accept or reject the feedback points that they receive, there is no guarantee that their revision of text, or writing performance will improve. Based on the findings from the think-aloud data in this study, it can be argued that the



cognitive operations and approaches employed by the treatment group participants may also help explain the better effects of the peer feedback intervention.

### 7.5.1 Cognitive operations and their efficacy

This study, based upon the data from 94 PF-triggered revision episodes collected from the six case study participants, identified nine cognitive operations that were employed in processing peer feedback. These include *referring to peer feedback*, *referring to text*, *retrieving L2 knowledge*, *justifying*, *resourcing*, *translating*, *monitoring*, *evaluating* and *generating changes/no change*. This study reveals a wider range of cognitive operations than those reported in previous peer feedback and teacher feedback studies (e.g., Fan & Xu, 2020; Han, 2016; Rajoo, 2009).

As shown in 5.3.1, apart from *referring to PF* and *generating changes/no change*, the two cognitive operations present in each PF-triggered revision episodes, *resourcing* was the most frequently employed cognitive operation. The six case study participants used online and paper dictionaries when writing their second drafts for Writing Task 6. M. Yang et al. (2006) and Villamil and de Guerrero (1996) also found their students used dictionaries and grammar books to confirm the validity of peer feedback and to solve difficulties when writing their second drafts. In this study, the teacher's emphasis on using reference resources seemed to be a major reason for the participants' frequent use of dictionaries. The teacher touched on this topic twice in the six focused training sessions, in the first session by exemplifying how to use dictionaries in feedback provision and in revision, and in the fifth session by introducing other reference resources, including corpus, scoring rubrics and exemplar compositions. Another reason might be that the participants believed they could get reliable and unambiguous answers from consulting dictionaries. In this study, *resourcing* appeared to be effective when the

participants consulted dictionaries to solve problems of spelling and collocation, as can be seen from the correct forms in their second drafts. However, in some instances when they searched for L2 words and phrases by using the Chinese-English translation function of online dictionaries, they did not obtain appropriate words as they had expected. Furthermore, when they copied some words or phrases mechanically without understanding their English meaning or use, they would be unlikely to recall them when writing on a new topic. Also, some participants may have been over-dependent on dictionaries. For instance, Liu, a case study participant, consulted dictionaries for the spelling of “technology” and the past tense for “put”, when she added some ideas as suggested by her peers. For problems that learners should be capable of solving by retrieving their L2 knowledge, repeatedly using dictionaries may affect their confidence in writing and retention of knowledge.

As well as *resourcing*, the case study participants were also found to employ the cognitive operations of *retrieving L2 knowledge*, *monitoring*, *justifying* and *evaluating* frequently in revisions triggered by peer feedback. Such findings lend support to and extend findings from previous teacher feedback and peer feedback studies. *Retrieving L2 knowledge* was similar to resorting to interlanguage knowledge, as noted in Villamil and de Guerrero (1996) and activating previous knowledge in Fan and Xu (2020) and Han (2016). Specifically, the case study participants in this study accessed L2 words, grammatical structures and writing conventions that they had learned previously. In most instances, the activated L2 knowledge resulted in valid changes in that the retrieved knowledge was mostly implicit knowledge, rather than explicit knowledge.

The use of the cognitive operation of *monitoring* by the case study participants in this study, similar to those in Fan and Xu (2020) and Han (2016), was evident in their

monitoring of changes they made in response to received feedback and of other related text segments which did not receive feedback. By constantly reminding themselves to check for possible errors through monitoring, students produced their second drafts with improved accuracy.

The cognitive operation of *justifying* was also reported by Han (2016), Qi & Lapkin (2001) and Rajoo (2009). In this study, the case study participants provided justifications for feedback they received and for changes they intended to make. Given that articulation of reasons for feedback was taken as evidence that learners understood teacher feedback (Qi & Lapkin, 2001), justification for a change should indicate that changes were made based upon careful consideration, but not random decisions.

The cognitive operation of *evaluating*, referring to students' checking a change against a standard, was not reported in previous peer feedback or teacher feedback studies. In this study, the case study participants checked some changes in terms of appropriateness or accuracy against the received feedback or their L2 intuition. When they judged the change to be inappropriate, at times they used other cognitive operations and revised in a different way. To some extent, the double-checking increased the likelihood that the changes in their second drafts were appropriate.

Of the nine cognitive operations identified in the present study, the least employed were *referring to text* and *translating*. A possible reason for the low frequency of *referring to text* is that the excerpts in which the case study participants read part of their texts to contextualize the received feedback were categorized as *referring to PF*, rather than *referring to text*. As defined in 3.7.3.2, only excerpts in which the case study participants read or paraphrased a text segment to obtain ideas was categorized as *referring to text*. For instance, when Liu responded to a feedback point asking her to add a transition

sentence between paragraphs, she referred to ideas in both paragraphs and incorporated them when adding a sentence in her second draft (see 5.3.1.2). Just as “the text produced so far” is an indispensable resource when writers compose (Flower & Hayes, 1981, p. 370), both the first drafts and on-going second drafts can be a resource in revision and students should be reminded to utilize them when needed.

The cognitive operation of *translating* was apparent when the case study participants in this study used their L1 to make meaning of L2 texts or to generate new L2 text segments, similar to those in Villamil and de Guerrero (1996). A possible reason for their infrequent use of L1 in revision may be the participants’ high L2 proficiency (see 3.6.1). Just as L2 learners have been reported to decrease L1 use in their L2 composing process as they develop in L2 proficiency (W. Wang & Wen, 2002), their dependence on L1 may also be reduced in the revision process. In addition, since some L1 use in the composing process has been found to have a negative impact on L2 text quality (Fujii, 2012; Kim, 2014), it might be desirable for teachers to guide L2 learners, particularly those with advanced L2 proficiency, in strategic L1 use in L2 revision tasks.

### 7.5.2 Two approaches and their efficacy

Findings from the case study show that the treatment group participants employed two approaches in processing peer feedback, the direct approach in which they only referred to peer feedback and generated changes or no change, and the elaborated approach in which they implemented other cognitive operation (s) in addition to *referring to PF* and *generating changes/no change*. The two approaches were similar to those identified by Storch and Wigglesworth’s (2010):

LREs [Language related episodes, similar to PF-triggered revision episodes in this study] that showed evidence of extensive engagement included episodes in

which learners offered suggestions and counter-suggestions, explanations, or any comments that showed evidence of meta-awareness of the feedback received...LREs that showed evidence of limited engagement included episodes in which one member of the pair simply read the feedback and the other merely acknowledged or repeated it. (p. 311)

In this study, the case study participants were found to employ the direct approach to tackle feedback points that targeted surface errors or that were thoroughly discussed with their peers. For surface errors, Storch and Wigglesworth (2010) argued that “limited or no overt engagement” would lead to uptake, and possibly retention (p.328). When students made changes that they had thoroughly and adequately discussed with their peers, the changes were more likely to be valid because jointly-constructed knowledge was more reliable than knowledge constructed individually.

The case study participants employed the elaborated approach when dealing with feedback on complex language problems, content and organization problems. Similarly, both Storch and Wigglesworth (2010) and Fan and Xu (2020) found their participants engaged extensively with feedback targeting complex language errors. For the processing of feedback points on content and organization problems, Fan and Xu (2020), however, reported a contradictory finding that their participants demonstrated low engagement with feedback on ideas and text organization. A possible reason for this discrepancy might be the attention to revision strategies, particularly in how to respond to feedback on content and organization, in peer feedback training. As the peer feedback training in Fan and Xu (2020) focused on “strategies of evaluating a composition”(p. 3), the participants may neither realize the importance of attending to feedback on content and organization, nor feel competent to deal with such feedback. However, in the present

study, the six focused training sessions giving priority to revision strategies, prepared the treatment group participants for strategies to use in responding to received peer feedback. This study also reveals that a critical attitude towards peer feedback was another factor leading to the case study participants' employment of the elaborated approach. As shown in 5.3.2.2, in response to feedback points on word choice or collocation, the participants in most instances consulted online dictionaries and possibly carried out other cognitive operations before incorporating feedback in their second drafts. The elaborated approach may have contributed to improvements in text revisions as it helped students achieve a better understanding of their problems and enabled them to make judicious decisions in revision. Moreover, it may have led to an improvement in writing performance as the multiple cognitive operations helped the writers internalize knowledge gained from solving the problems.

## **7.6 Issues in peer feedback intervention**

Despite the fact that peer feedback intervention was successful and may have contributed to improved writing performance, there were still some issues identified by the treatment group participants in the questionnaires. Such issues are equally worthy of attention in that they will inform future studies.

Firstly, the treatment group participants raised two issues pertaining to paired discussions: time constraints and repetition of written comments. In this study, three treatment group participants indicated that thirty minutes was not sufficient to discuss two compositions and accompanying feedback points. The participants in W. Wang (2014) similarly reported that insufficient discussion time prevented them from working out solutions to problems in their compositions. To address the tension between tight schedules during class hours and students' need for more time for discussions, previous

studies have proposed some suggestions. W. Wang (2014) suggested that students be required to provide written feedback before class and class time be used only for negotiation, as implemented in this study. Q. Zhu (2018) suggested that teachers narrow the scope of discussion to allow “in-depth dialogue to clarify, negotiate and explore meanings” (p. 167), or incorporate a flipped class to offer some instruction in peer feedback as out-of-classroom activities, reserving the class hours for discussions. It is suggested that future peer feedback studies include these practices to improve the efficacy of discussions. The problem of repetition was reported by five participants as a factor affecting their perceived helpfulness of paired discussions. When peer reviewers only repeated their written comments, they obviously displayed passivity and unwillingness to collaborate in interactions, which was “detrimental to the functioning of peer response/review, leading to little response to peer comments in subsequent revisions” (Min, 2008, p. 287). Although stances and attitudes in peer interactions were touched upon in the general training session in this study (see 2.3.3), such comments suggest that some key concepts and guidelines for peer feedback activities might need to be revisited during the intervention period.

Secondly, it is somewhat surprising that as many as 13 treatment group participants raised the problem of using PF forms. The participants in this study, just as those in Gielen et al (2010) who expressed their strong dislike for the “paperwork [PF form]” (p. 159), considered completing PF forms unhelpful and time-consuming and voiced their dissatisfaction with some elements in PF forms. Such perceptions need to be challenged; otherwise, they may affect the participants’ appreciation of the value of PF forms in facilitating feedback and providing guidance on writing (see 6.3.4). Participants, who expressed their preference for writing comments on their peer’s first draft suggested a practical solution: Teachers may design PF forms as a checklist to ensure that the forms

fulfil their function of facilitating feedback, while allowing students to decide how to provide written feedback, as in-text comments, end-of-text comments, or other possible forms. In addition, if students could be involved in designing PF forms and updating their content regularly, negative perceptions arising from their dissatisfaction with PF forms may be counteracted.

Thirdly, although various forms of teacher support, the general training session, focused training sessions and PF forms, were incorporated in this study, five treatment group participants still raised their concerns about the quality of peer feedback. Students' reservations about the quality of peer feedback have been consistently documented in peer feedback literature (Lee, 2015; Mangelsdorf, 1992; Nelson & Carson, 1998), and researchers have proposed several solutions to tackle this issue, the most common of which is using teacher feedback in conjunction with peer feedback. Early L2 peer feedback studies (e.g., Tsui & Ng, 2000; M. Yang et al., 2006) used peer feedback in combination with teacher feedback to allow students to receive different forms of feedback at different stages. Later studies (e.g., W. Wang, 2014; Zhao, 2014) have encouraged provision of teacher feedback on peer feedback. The two practices, however, cannot address the students' concern about peer feedback quality. When teachers identify problems that that peers fail to notice or rectify, students' dependence on teacher feedback may increase while their trust in peer feedback is likely to decrease. To mitigate students' concerns about peer feedback quality, incorporating what has been termed, intra-feedback, into peer feedback practice can be a viable practice, as Lee (2015) posited:

Intra-feedback was carried out in a reviewers' meeting where two reviewers of the same essays first presented all of their written comments on the essays



(offered individually prior to the reviewers' meeting) to each other. The reviewers then discussed with each other any potential discrepancies or uncertainties in their comments, offered feedback on each other's feedback performance, and, if necessary, revised their own comments before presenting them to the writers of the essays. (Lee, 2015, p. 3)

### **7.7 Further development of Hayes' (1996) revision model**

Based on the cognitive operations identified in this study and previous studies on feedback, an adapted Hayes' (1996) cognitive process model of revision is proposed, as shown in Figure 7.1.

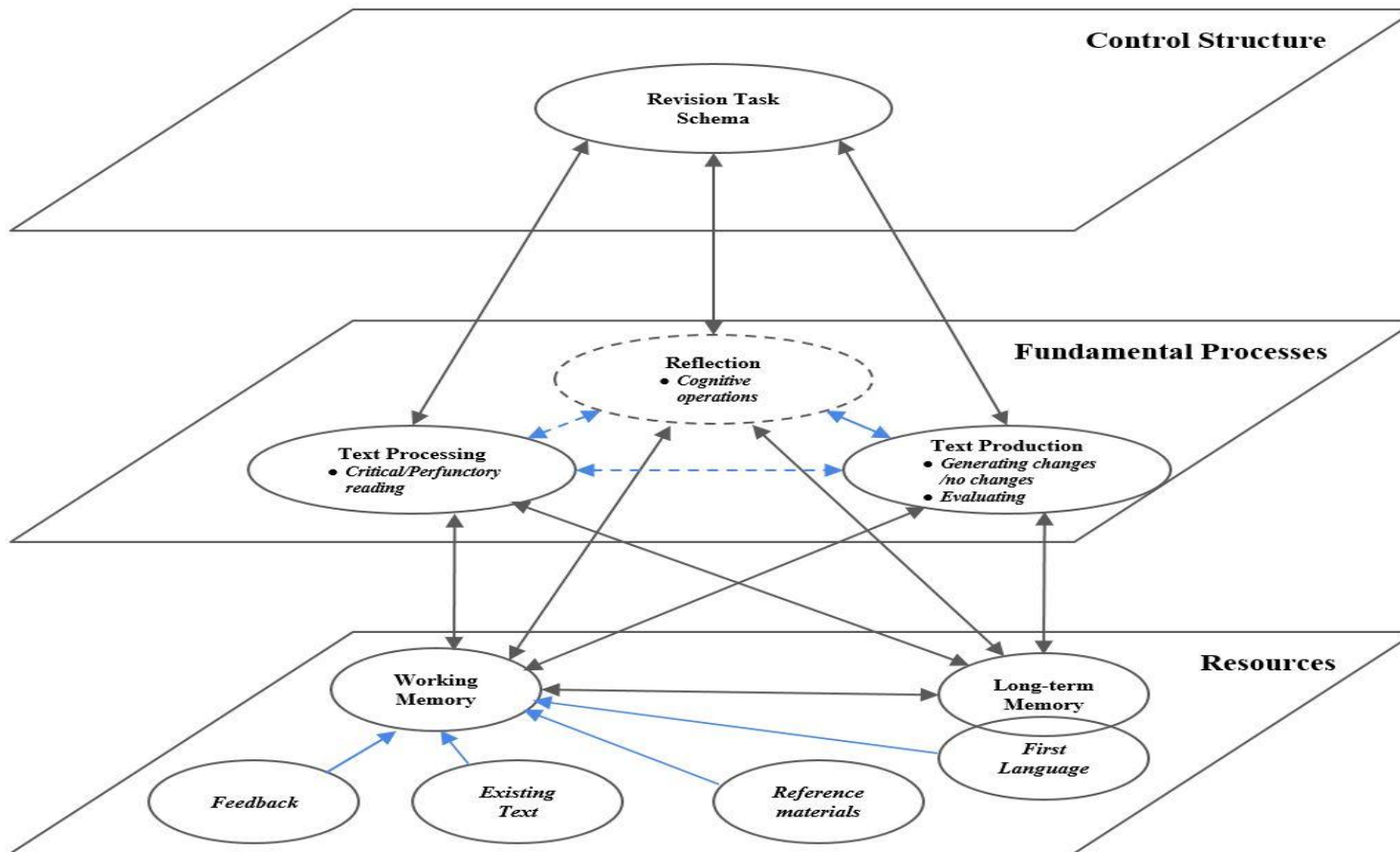


Figure 7.1 The adapted Hayes' (1996) cognitive process model of revision

*Note. Cognitive operations include referring to text, retrieving L2 knowledge, justifying, resourcing, translating, monitoring, memorizing, making connections and reflecting*

Each revision attempt starts with text processing in which the writers read their texts, and/or feedback they have received, their reading of which can be critical or perfunctory. When EFL writers read their texts and/or feedback to detect problems, they read critically (Dost, 2019; Rajoo, 2009). There are, however, some instances when they read some feedback points with only partial or even no understanding and hastily move to the stage of text production (Han, 2016; Qi & Lapkin, 2001). In such instances, they read perfunctorily. Although critical reading is optimum for revision, perfunctory reading should also be included in order to objectively reflect how EFL writers go through the text processing stage.

The reflection process occurs after text processing. The EFL writers, compared with their L1 counterparts, have a rich repertoire of cognitive operations to choose from, including *referring to text*, *retrieving L2 knowledge*, *justifying* (providing reasons for a received feedback point or for change(s) to be made), *resourcing* (consulting reference materials), *translating*, *monitoring*, *memorizing* [feedback and solutions], *making connections* [between feedback points], and *reflecting* (making judgement about the validity of a certain feedback point). When dealing with a specific problem, they may employ multiple cognitive operations or conduct certain cognitive operation(s) repeatedly. Since both this study and some previous ones (Fan & Xu, 2020; Han, 2016; Storch & Wigglesworth, 2010) have found that the reflection process was only present when the writers responded to feedback points on complex issues (word choice, sentence structure, content, or organization problems), the circle line of the reflection process is dashed in the adapted model, indicating reflection as an optional process.

The last fundamental process is text production, with two cognitive operations, *making changes/no change*, a core cognitive operation in the text production process, and

*evaluating*. Based on the findings from this study that some participants checked the change(s) they made in terms of appropriateness, *evaluating* is added to the text production process in the adapted model.

Also, in the adapted model, the three fundamental processes interact with each other as indicated by the double-headed arrows in blue. A solid double-headed arrow is drawn between the reflection and text production processes based on the findings from this study that in some PF-triggered revision episodes, the cognitive operation of generating changes occurred multiple times with each preceded by some cognitive operations in the reflection process. For the potential interactions between text-processing and reflection and between text-processing and text-production, dashed double-headed arrows are used.

The third component of the adapted model is resources. While the two components, working memory and long term memory and their interactions with other components in Hayes' (1996) revision model are retained, another four resources that EFL writers have access to in revision are added; these include feedback (from peers, from teachers, and from other sources), existing text (their first draft and their on-going second draft), reference materials (e.g., dictionaries and grammar books), and their first language. The newly added resources are indicated by italics in the adapted model. In addition, the newly added resources and working memory are connected with single-headed arrows in that these resources are only temporarily activated for revision purpose. Also of note is the overlap between the long-term memory and the first language, indicating some knowledge (topic knowledge and rhetorical knowledge) in the long-term memory may be L1-mediated.

## **7.8 Chapter summary**

This chapter has discussed the key findings of this study within the wider context of the theoretical framework and key empirical studies. Some of the key findings of this study agree with previous studies, whereas others do not. Both the peer feedback intervention and the participants' efforts in utilizing peer feedback in revision may have contributed to the positive effects of peer feedback on text revisions and writing performance. Finally, some adaptations are made to Hayes' (1996) revision model on the basis of findings from this study and previous feedback studies.

# **Chapter 8 Conclusions, Implications, and Recommendations**

## **8.1 Chapter overview**

This chapter first summarizes the major findings of this study. It then explains the theoretical and methodological contributions of this study as well as its pedagogical implications. Finally, it discusses the limitations of this study and proposes recommendations for future research.

## **8.2 Summary of major findings**

This study implemented a rigorous peer feedback intervention and examined its effects on Chinese EFL learners' text revisions and writing performance as a whole, drawing upon various data sources ranging from product data, process data to perception data. Major findings are presented as follows.

### **8.2.1 Effects of peer feedback on text revisions**

The peer feedback intervention implemented in this study was more effective than collective feedback in helping Chinese EFL learners make revisions, as seen in the greater gains in the overall score made by the treatment group participants between the two drafts than the comparison group. The two groups performed differently in text revisions, in terms of content, organization and language use, as noted below.

- (1) The treatment group made greater gains between drafts in content and organization scores, between drafts, than the comparison group.

- (2) The treatment group and the comparison group had different outcomes in language use. The treatment group and the comparison group made similar improvements in accuracy; only the treatment group made a significant improvement in syntactic complexity; neither group made a significant improvement in lexical complexity or fluency.

### 8.2.2 Effects of peer feedback on writing performance

The peer feedback intervention was more effective than collective feedback in helping students' improvements in writing performance, as seen from the changes in overall scores for the two groups. Both the treatment group and the comparison group improved significantly in overall scores in the post-test and delayed post-test, but the treatment group outperformed the comparison group in the post-test and delayed post-test. Their performance in written content, organization and language use are presented as follows.

- (1) With content quality, only the treatment group made a significant improvement in the post-test, which was retained in the delayed post-test. There was no significant difference between the two groups in the post- and delayed post-tests.
- (2) Organization quality improved significantly only for the treatment group and they retained their improvement in the delayed post-test. The treatment group significantly outperformed the comparison group in both the post- and delayed post-tests.
- (3) With language use, both the treatment group and the comparison group made similar improvements in accuracy in the post-test, but only the treatment group retained their improvement in the delayed post-test. Only the treatment group made a significant improvement in the post-test in fluency, but the effect was

not retained in the delayed post-test. There was no significant change for either group in lexical complexity and syntactic complexity in the post-test or delayed post-test.

### **8.2.3 Processing of peer feedback in revision**

The findings from the think-aloud data reveal that the six case study participants employed nine cognitive operations in processing peer feedback in revision: *referring to PF*, *referring to text*, *retrieving L2 knowledge*, *justifying*, *resourcing*, *translating*, *monitoring*, *evaluating* and *generating changes/no change*. In addition, the two approaches in processing peer feedback, the direct approach and elaborated approach, as well as factors influencing the case study participants' use of each approach are presented.

### **8.2.4 Perceptions of peer feedback intervention**

The findings from the open-ended questionnaires and interviews showed the treatment group participants' strong belief about their improvements in written content, organization and accuracy as well as their positive attitudes towards the peer feedback intervention. The findings suggest that three issues might affect effectiveness of the intervention: concerns for peer feedback quality, time constraints for paired discussions in class and compulsory use of PF forms.

## **8.3 Contributions and implications**

This study examined the effects of a rigorous peer feedback intervention and explained its advantage in helping Chinese EFL learners improve text revisions and writing performance. Its contributions to the existing literature and its pedagogical implications are presented in the sub-sections below.



### 8.3.1 **Theoretical contributions**

Firstly, the findings from this study add strong evidence to the long-standing debate on the effects of peer feedback. It can be argued that improvements in text revisions and writing performance demonstrated by Chinese EFL learners were the outcome of the peer feedback intervention implemented in this study, informed by several key concepts in sociocultural theory, including mediation, scaffolding and Zone of Proximal Development (ZPD). Evidence was presented of the long-term positive impact of peer feedback on writing performance, twelve weeks after the intervention. Existing peer feedback studies have rarely considered its long-term effects on writing performance.

Secondly, findings about the nine cognitive operations and the two approaches which EFL learners employed in processing peer feedback in revision provide a new perspective on Hayes' (1996) revision model. The adapted Hayes' (1996) revision model, developed in this study, incorporates feedback, a critical element in revision, as well as other possible resources to which EFL learners have access. The cognitive operations in the adapted revision model, identified in the present study and reported in existing feedback studies, suggest that revision is a process in which writers are selective in using cognitive operations. A strength of the adapted model, compared with Hayes' (1996) revision model, is that it gives consideration to feedback and other resources in revision, and most importantly, the needs of EFL writers.

### 8.3.2 **Methodological contributions**

Methodologically, this study adopted a mixed-method approach and included multiple instruments to collect data. While the quasi-experimental study included the treatment group and comparison group to compare their performance in text revisions and writing performance over an extended period, the perception data gathered from questionnaires

and interviews and the process data collected from the case study helped to cross-check and explain the quantitative findings. In addition, this study, with clear explanations to its procedures and presentation of supporting documents in the Appendices, can be replicated in other EFL contexts.

In addition, this study featured a long intervention period with multiple feedback iterations. Some previous peer feedback studies tended to be one-off, or involved only a few writing tasks, and have thus lacked theoretical and pedagogical validity (Storch, 2010). By extending the duration of the intervention to one semester and distributing the six writing tasks evenly within the semester, this study allowed multiple peer feedback practices for the participants and sufficient time for the progress in text revisions and in writing development to be assessed.

Moreover, this study highlighted the combined use of writing scores and CAF indexes in evaluating writing quality. The multiple measures used in this study helped capture dynamic changes in writing associated with the intervention.

### **8.3.3 Pedagogical implications**

The findings of this study have pedagogical implications. Firstly, on-going support from EFL writing teachers is crucial for effectiveness of an intervention. In this study, teacher support was provided not only through the general training, but also through PF forms and several focused training sessions, which made peer feedback task manageable and addressed challenges arising from actual peer feedback activities. While findings from this study indicate that learners appreciated the value of different forms of teacher support, they also alert EFL writing teachers to some issues in peer feedback activities, such as how to use PF forms and what to include in PF forms. When EFL writing teachers offer support, they need to bear in mind that their support should positively

impact, rather than interfere with, learners' performance in and perceptions of peer feedback activities.

Secondly, EFL writing teachers need to address the dilemma about time constraints for peer feedback activities in class. While peer feedback activities, undeniably, take up a substantial amount of class hours, the time investment is worthwhile, as evident in the participants' comments about the contribution of focused peer feedback training sessions and paired discussions and their demand for more time for paired discussions. Caught between tight teaching schedules during class hours and the need for more time for peer feedback activities, EFL writing teachers should devise ways to maximize the limited in-class time, without either sacrificing the intervention or other curricular activities.

Thirdly, the coding scheme for cognitive operations, the consequent two approaches in processing peer feedback in revision, and the adapted revision model could be used to train learners explicitly on using peer feedback in revision. Such training will help learners develop an understanding about what resources are available and how specific cognitive operations and approaches can be used effectively. Familiarity with available resources, cognitive operations and approaches would enable students to gain optimum benefits from peer feedback.

#### **8.4 Limitations and recommendations for future research**

Although this study was longitudinal and included quantitative data and qualitative data, some limitations remain. Firstly, as all the writing tasks in this study were expository topics, the effects of the intervention need to be interpreted with caution, as the results may not generalize to other genres. Further studies could incorporate other genres and examine whether effects of peer feedback vary according to genres.

Secondly, the findings that there was no change in some measures for language use may suggest that these measures were inappropriate for the investigation into efficacy of peer feedback. Given that there are a variety of measures for each dimension of language use, and many empirical studies have incorporated CAF indexes, researchers could consider using multiple measures for each dimension and make judicious decisions about the measures.

Thirdly, revisions that were not linked to peer feedback were not explored in the present study. Little is known about what triggers such revisions, as the think-aloud method elicits real-time data and discourages participants from explaining or justifying what they are doing. Although some previous studies have used text analysis of peer feedback and changes between drafts to infer the link between peer feedback and revisions, future studies could include stimulated recall interviews to cross-check findings from text analysis and provide reliable evidence for relations between peer feedback and revisions.

Finally, the findings about how learners process peer feedback in revision are only preliminary because the case study involved only six participants and one writing task. Future studies could extend the scope of this study by involving a larger sample of participants, exploring how EFL learners with different writing proficiencies process peer feedback in revision, or comparing how EFL learners process peer feedback and other forms of feedback in revision.

## **8.5 Conclusion**

Based on the findings of this study, the conclusion is that peer feedback can be incorporated into EFL writing instruction to help EFL students improve their text revisions and writing ability. Compared with collective feedback, the peer feedback intervention implemented in this study was more effective in helping Chinese EFL

learners improve in text revisions as well as in writing performance. Importantly, the positive impact of peer feedback on writing performance maintained twelve weeks after the intervention. From this study, teachers and researchers working with EFL learners may gain valuable insights as regards the design and implementation of peer feedback activities in EFL writing instruction.

## References

- Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research*, 54(2), 139–172.
- Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht, The Netherlands: Kluwer.
- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465–483.
- Allal, L., & Chanquoy, L. (2004). Introduction: Revision revisited. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 1–7). New York, NY: Kluwer.
- Anson, I. G., & Anson, C. M. (2017). Assessing peer and instructor response to writing: A corpus analysis from an expert survey. *Assessing Writing*, 33, 12–24.
- Atkinson, D. (2003a). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing*, 12(1), 3–15.
- Atkinson, D. (2003b). Writing and culture in the post-process era. *Journal of Second Language Writing*, 12(1), 49–63.
- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67–81.

- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research, 12*(3), 409–431.
- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal, 63*(3), 204–211.
- Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics, 31*(2), 193–214.
- Bui, G., & Kong, A. (2019). Metacognitive instruction for peer review interaction in L2 writing. *Journal of Writing Research, 11*(2), 357–392.
- Butterfield, E. C., Hacker, D. J., & Albertson, L. R. (1996). Environmental, cognitive, and metacognitive influences on text revision: Assessing the evidence. *Educational Psychology Review, 8*(3), 239–297.
- Cai, J. (2017). Debates around the orientation of TEFL in Chinese tertiary education. In H. Reinders, D. Nunan, & B. Zou (Eds.), *Innovation in Language learning and teaching* (pp. 115–153). London, England: Palgrave Macmillan.
- Cao, Z., Yu, S., & Huang, J. (2019). A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Studies in Educational Evaluation, 63*, 102–112.
- Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed

- first-year L2 writing. *Journal of Second Language Writing*, 44, 51–62.
- Casanave, C. P. (2003). Looking ahead to more sociopolitically-oriented case study research in L2 writing scholarship (But should it be called “post-process”?). *Journal of Second Language Writing*, 12(1), 85–102.
- Chang, C. Y. (2016). Two decades of research in L2 peer review. *Journal of Writing Research*, 8(1), 81–117.
- Chanquoy, L. (2009). Revision processes. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The SAGE handbook of writing development* (pp. 80–97). London, England: SAGE.
- Charter, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education Journal*, 12(2), 68–82.
- Cheng, J., & Wei, L. (2021). Individual agency and changing language education policy in China: reactions to the new “Guidelines on College English Teaching.” *Current Issues in Language Planning*, 22, 1–19.
- Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328–338.
- Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103(1), 73–84.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers and Education*, 48(3), 409–426.



- Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication, 23*(3), 260–294.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology, 98*(4), 891–901.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Connor, U., & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing, 3*(3), 257–276.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5<sup>th</sup> ed.). Thousand Oaks, CA: SAGE.
- Crinon, J. (2012). The dynamics of writing and peer review at primary school. *Journal of Writing Research, 4*(2), 121–155.
- 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.
- Cumming, A., Busch, M., & Zhou, A. (2002). Investigating learners' goals in the context of adult second-language writing. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 189–208). Dordrecht, The Netherlands: Kluwer.

- de Guerrero, M. C. M., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *The Modern Language Journal*, 78(4), 484–496.
- 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.
- deBoer, F. (2014). Evaluating the comparability of two measures of lexical diversity. *System*, 47, 139–145.
- Diab, N. M. (2010). Effects of peer- versus self-editing on students' revision of language errors in revised drafts. *System*, 38(1), 85–95.
- 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.
- Diab, N. M. (2016). A comparison of peer, teacher and self-feedback on the reduction of language errors in student essays. *System*, 57, 55–65.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. New York, NY: Oxford University Press.
- Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research*. New York, NY: Routledge.
- Dost, G. T. (2019). A revision model in writing for novice writers with a focus on audience and feedback. *Journal of Contemporary Educational Research*, 3(1), 6–

- Dressler, R., Chu, M., Crossman, K., & Hilman, B. (2019). Quantity and quality of uptake: Examining surface and meaning-level feedback provided by peers and an instructor in a graduate research course. *Assessing Writing*, 39, 14–24.
- Duff, P. A. (2020). Case study research: Making language learning complexities visible. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 145–153). New York, NY: Routledge.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, England: Oxford University Press.
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford, England: Oxford University Press.
- Emig, J. (1971). *The composing processes of twelfth graders*. Urbana, IL: National Council of Teachers of English.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2<sup>nd</sup> ed.). Cambridge, MA: MIT Press.
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32(4), 400–414.
- Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing*, 50, 100775.
- Ferris, D., & Eckstein, G. (2020). Language matters: Examining the language-related needs and wants of writers in a first-year university writing course. *Journal of*

- Writing Research*, 12(2), 321–364.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- Fujii, Y. (2012). The effects of L1 on L2 writing and translation: A case study. *Journal of Modern Languages*, 22(1), 32–44.
- Gai, S., & Zhou, X. (2013). 基于动态系统理论的同伴反馈实证研究——社会文化观视角 [A dynamic approach to the empirical study on EFL peer-group prior-evaluation: From a sociocultural perspective]. *Foreign Language and Their Teaching*, 269, 36–40.
- Gallien, T. & Oomen-Early, J. (2008). Personalized versus collective instructor feedback in the online classroom: Does types of feedback affect student satisfaction, academic performance and perceived connected ness with the instructor? *International JI. on E-learning*, 7(3), 463-476.
- Gao, Y., Schunn, C. D. D., & Yu, Q. (2019). The alignment of written peer feedback with draft problems and its impact on revision in peer assessment. *Assessment & Evaluation in Higher Education*, 44(2), 294–308.
- Gardner, S., & Nesi, H. (2013). A classification of genre families in university student writing. *Applied Linguistics*, 34(1), 25–52.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.
- Geng, H. (2017). *The effects of written feedback and revision-focus direction on student*

- revision and writing improvement*. Unpublished PhD thesis, University of Auckland, Auckland, New Zealand.
- Gielen, S., Tops, L., Dochy, F., Onghena, P., & Smeets, S. (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, 36(1), 143–162.
- Gil, J. (2016). English language education policies in the People’s Republic of China. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 49–90). Cham, Switzerland: Springer.
- Graner, M. H. (1987). Revision workshops: An alternative to peer editing groups. *The English Journal*, 76(3), 40–45.
- Hamp-Lyons, L. (1990). Second language writing: assessment issues. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 69–87). New York, NY: Cambridge University Press.
- Han, Y. (2016). *Learner engagement with teacher written corrective feedback in Chinese tertiary-level EFL classrooms: A sociocognitive perspective*. Unpublished PhD thesis, University of Hong Kong, HKSAR, China.
- Haswell, R. H. (2005). NCTE/CCCC’s recent war on scholarship. *Written Communication*, 22(2), 198–223.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Erlbaum.

- Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J. F., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in psycholinguistics* (Vol. 2, pp. 176–240). Cambridge, England: Cambridge University Press.
- He, F., & Teng, M. F. (2019). Language tug-of-war: When English literacy education encounters the National Matriculation English Test policy in mainland China. In B. L. Reynolds & M. F. Teng (Eds.), *English literacy instruction for Chinese speakers* (pp. 229–316). Singapore: Palgrave Macmillan.
- Housen, A., Kuiken, F., & Vedder, I. (2012) *Dimensions of L2 performance and proficiency complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins Publishing.
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), 321–342.
- 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.
- Huang, Y. (2020). *A process-genre approach to teaching writing: Effects on Chinese EFL learners' argumentative writing quality, genre knowledge and metacognitive strategies*. Unpublished PhD thesis, University of Auckland, Auckland, New Zealand.
- Huang, Y., & Zhang, L. J. (2020). Does a process-genre approach help improve students' argumentative writing in English as a Foreign Language? Findings from an intervention study. *Reading & Writing Quarterly*, 36(4), 339–364.

- Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English*, 43(7), 732–739.
- Iwaniec, J. (2020). Questionnaires: Implications for the effective implementation. In *The Routledge handbook of research methods in applied linguistics* (pp. 325–335). New York, NY: Routledge.
- Jacobs, H. L., Zingraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- 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), 12–39.
- Karegianes, M. L., Pascarella, E. T., & Pflaum, S. W. (1980). The effects of peer editing on the writing proficiency of low-achieving tenth grade students. *The Journal of Educational Research*, 73(4), 203–207.
- Karpov, Y. V., & Haywood, H. C. (1998). Two ways to elaborate Vygotsky's concept of mediation. *American Psychologist*, 53(1), 27–36.
- Katstra, J., Tollefson, N., & Gilbert, E. (1987). The effects of peer evaluation on attitude toward writing and writing fluency of ninth grade students. *The Journal of Educational Research*, 80(3), 168–172.
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, 39(3), 387–406.
- Kim, Y. (2014). The use of L1 as a writing strategy in L2 writing tasks. *GEMA Online®*

- Journal of Language Studies*, 14(3), 33–50.
- Lam, R. (2010). A peer review training workshop: Coaching students to give and evaluate peer feedback. *TESL Canada Journal*, 27(2), 114–127.
- Lam, R. (2013). The relationship between assessment types and text revision. *ELT Journal*, 67(4), 446–458.
- Lantolf, J. P. (2000). Second language learning as a mediated process. *Language Teaching*, 33(2), 79–96.
- Lee, M.-K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, 55, 1–10.
- Li, H. H., Zhang, L. J., & Parr, J. M. (2020). Small-group student talk before individual writing in tertiary English writing classrooms in China: Nature and insights. *Frontiers in Psychology*, 11, 1–14.
- Li, H., Zhong, Q., & Suen, H. K. (2012). Students' perceptions of the impact of the College English Test. *Language Testing in Asia*, 2(3), 77–94.
- Lindgren, E. (2005). The uptake of peer-based intervention in the writing classroom. In G. Rijlaarsdam, H. Bergh, & M. Couzijn (Eds.), *Effective learning and teaching of writing: A handbook of writing in education* (2<sup>nd</sup> ed., pp. 259–274). New York, NY: Kluwer.
- Liou, H.-C., & Peng, Z.-Y. (2009). Training effects on computer-mediated peer review. *System*, 37(3), 514–525.



- Llosa, L., Beck, S. W., & Zhao, C. G. (2011). An investigation of academic writing in secondary schools to inform the development of diagnostic classroom assessments. *Assessing Writing, 16*(4), 256–273.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning, 45*(4), 605–655.
- Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207–228.
- 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.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? *ELT Journal, 46*(3), 274–284.
- Mangelsdorf, K., & Schlumberger, A. (1992). ESL student response stances in a peer-review task. *Journal of Second Language Writing, 1*(3), 235–254.
- Matsuda, P. K. (2003). Process and post-process: A discursive history. *Journal of Second Language Writing, 12*(1), 65–83.
- McGroarty, M. E., & Zhu, W. (1997). Triangulation in classroom research: A study of peer revision. *Language Learning, 47*(1), 1–43.
- Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28*(4), 745–769.
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and*

- implementation* (4<sup>th</sup> ed.). San Francisco, CA: Wiley.
- Min, H.-T. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293–308.
- Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118–141.
- Min, H.-T. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes*, 27(3), 285–305.
- Muijs, D. (2010). *Doing quantitative research with SPSS*. London, England: SAGE.
- Nelson, G., & Carson, J. (2006). Cultural issues in peer response: Revisiting “culture”. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 42–59). New York, NY: Cambridge University Press.
- Nelson, G. L., & Carson, J. G. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing*, 7(2), 113–131.
- Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27(1), 135–141.
- Ng, C. H., & Cheung, Y. L. (2017). Innovations in writing instruction in China: Metasynthesis of qualitative research for the period 2005-2016. In H. Reinders, D. Nunan, & B. Zou (Eds.), *Innovation in language learning and teaching* (pp. 63–88). Singapore: Palgrave Macmillan.
- Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar.

- In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51–78). New York, NY: Oxford University Press.
- Olson, V. L. B. (1990). The revising processes of sixth-grade writers with and without peer feedback. *Journal of Educational Research*, 84(1), 22–29.
- O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. New York, NY: Cambridge University Press.
- Ortega, L. (2015). Syntactic complexity in L2 writing; Progress and expansion. *Journal of Second Language Writing*, 29, 82-94.
- Oxford, R. L. (2016). *Teaching and researching language learning strategies: Self-regulation in context*. New York, NY: Routledge.
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265–289.
- Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.
- Poverjuc, O., Brooks, V., & Wray, D. (2012). Using peer feedback in a Master's programme: A multiple case study. *Teaching in Higher Education*, 17(4), 465–477.
- Prater, D., & Bermudez, A. (1993). Using peer response groups with limited English proficient writers. *Bilingual Research Journal*, 17(1–2), 99–116.
- 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.
- 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.
- Rajoo, M. (2009). *Feedback and revision: A protocol analysis*. Unpublished PhD thesis, University of Otago, Dunedin, New Zealand.
- Rastgou, A., Storch, N., & Knoch, U. (2020). The effect of sustained teacher feedback on CAF, content and organization in EFL Writing. *Iranian Journal of Language Teaching Research*, 8(2), 41–61.
- Ruegg, R. (2015a). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, 73–82.
- Ruegg, R. (2015b). Differences in the uptake of peer and teacher feedback. *REL C Journal*, 46(2), 131-145.
- Ruegg, R. (2017). Learner revision practices and perceptions of peer and teacher feedback. *Writing and Pedagogy*, 9(2), 275–300.
- Ruegg, R., & Lida, P. (2020). A meta-investigation of the use of the labels 'longitudinal' and 'long-term' in studies of feedback on writing. *Journal of Second Language Teaching and Research*, 8, 1-22.
- Sanchez, H. S., & Grimshaw, T. (2020). Stimulated recall. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 312–323). New York, NY: Routledge.
- Sengupta, S. (1998) Peer evaluation: 'I am not the teacher'. *ELT Journal*, 52(1), 19-26.
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An

- exploratory study. *Journal of Second Language Writing*, 9(3), 259–291.
- Sasaki, M., Mizumoto, A., & Murakami, A. (2018). Developmental trajectories in L2 writing strategy use: A self-regulation perspective. *Modern Language Journal*, 102(2), 292–309.
- Scardamalia, M., & Bereiter, C. (1983). The development of evaluative, diagnostic, and remedial capabilities in children's composing. In M. Martlew (Ed.), *The psychology of written language: Developmental and educational perspectives* (pp. 67–95). New York, NY: Wiley.
- Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Erlbaum.
- Shang, H. (2019). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*, 27, 1–13.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556–569.
- Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 22(3), 286–306.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1), 38–62.

- Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing, 1*(3), 217–233.
- Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System, 27*(3), 363–374.
- Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies, 10*(2), 29–46.
- Storch, N., & Wigglesworth, G. (2010). Learners processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition, 32*(2), 303–334.
- Suzuki, M. (2008). Japanese learners' self revisions and peer revisions of their written compositions in English. *TESOL Quarterly, 42*(2), 209–233.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). New York, NY: Oxford University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320–337.
- Teng, L. S., & Zhang, L. J. (2016). Fostering strategic learning: The development and validation of the Writing Strategies for Motivational Regulation Questionnaire (WSMRQ). *The Asia-Pacific Education Researcher, 25*(1), 123–134.
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language

- classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48, 100701.
- Thirakunkovit, P. (2019). A meta-analysis of effectiveness of teacher and peer feedback: Implications for writing instructions and research. *The Asian EFL Journal*, 21(1), 140–177.
- Trimbur, J. (1989). Consensus and difference in collaborative learning. *College Composition and Communication*, 51(6), 602–616.
- Tsui, A. B., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296.
- van den Berg, I., Admiraal, W., & Pilot, A. (2006). Peer assessment in university teaching: Evaluating seven course designs. *Assessment & Evaluation in Higher Education*, 31(1), 19–36.
- 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.
- Villamil, O. S., & de Guerrero, M. C. M. (1998). Assessing the impact of peer revision on L2 Writing. *Applied Linguistics*, 19(4), 491–514.
- Villamil, O. S., & Guerrero, M. C. M. (2019). Sociocultural theory: A framework for understanding the socio-cognitive dimensions of peer feedback. In K. Hyland & F.

- Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (2<sup>nd</sup> ed., pp. 26–44). Cambridge, England: Cambridge University Press.
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, 96(4), 576–598.
- Vygotsky, L. V. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. V. (1989). Concrete human psychology. *Soviet Psychology*, 27(2), 53–77.
- Wang, S. (2016). 《大学英语教学指南》要点解读 [An interpretation of the Guidelines on College English Teaching]. *Foreign Language World*, 174(3), 2–10.
- Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, 19, 80–96.
- Wang, W., & Gao, X. (2008). English language education in China: A review of selected research. *Journal of Multilingual and Multicultural Development*, 29(5), 380–399.
- Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, 11(3), 225–246.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge, England: Cambridge University Press.
- Wigglesworth, G., & Storch, N. (2012). Feedback and writing development through collaboration: A socio-cultural approach. In R. M. Manchón (Ed.), *L2 writing*



- development: Multiple perspectives* (pp. 69–100). Berlin, Germany: De Gruyter Mouton.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Honolulu, HI: University of Hawaii Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *The Journal of Child Psychology and Psychiatry*, 17(2), 89–100.
- Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, 39(4), 510–522.
- Yang, C., Zhang, L. J., & Parr, J. M. (2020). The reactivity of think-alouds in writing research: quantitative and qualitative evidence from writing in English as a foreign language. *Reading and Writing*, 33(2), 451–483.
- Yang, L., Yang, M., & Zhang, Y. (2013). 我国英语写作教学三种反馈方式的对比研究 [A comparative study of feedback styles in the teaching of EFL writing]. *Foreign Language Education*, 34(3), 63–67.
- 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.
- Yang, Y., & Meng, W. (2013). The effects of online feedback training on students' text revision. *Language Learning & Technology*, 17(2), 220–238.
- Yin, R. K. (1994). *Case study research: Design and methods* (2<sup>nd</sup> ed.). London, England:

SAGE.

- Yin, R. K. (2015). *Qualitative research from start to finish*. New York, NY: Guilford.
- You, X. (2004). “The choice made from no choice”: English writing instruction in a Chinese university. *Journal of Second Language Writing, 13*(2), 97–110.
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching, 49*(4), 461–493.
- Yu, S., Zhang, Y., Zheng, Y., Yuan, K., & Zhang, L. (2019). Understanding student engagement with peer feedback on master’s theses: A Macau study. *Assessment & Evaluation in Higher Education, 44*(1), 1–16.
- Zhang, L. J. (2016). Reflections on the pedagogical imports of western practices for professionalizing ESL/EFL writing and writing-teacher education. *Australian Review of Applied Linguistics, 39*(3), 203–232.
- Zhang, L. J., & Cheng, X. (2020). 我国近十年同伴反馈研究:回顾与展望 [A synthetic review of a decade of peer feedback research in China: Looking back and looking forward]. *Journal of Xi’an International Studies University, 28*(1), 48–56.
- Zhang, L. J., & Cheng, X. (2021). Examining the effects of comprehensive written corrective feedback on L2 EAP students’ linguistic performance: A mixed-methods study. *Journal of English for Academic Purposes, 54*, 101043.
- Zhang, L. J., Liu, X., & Liu, B. (2018). 运用分析性评估标准引导大学英语写作课堂动态评估与反馈 [The effect of implementing analytic scoring as dynamic assessment on student writing in a College English writing class]. *Fudan Forum*

*on Foreign Languages and Literature*, 11(1), 3–10.

- Zhang, L. J., & Zhang, D. (2020). Think-aloud protocols. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 302–311). New York, NY: Routledge.
- Zhang, X., & McEneaney, J. E. (2020). What is the influence of peer feedback and author response on Chinese university students' English writing performance? *Reading Research Quarterly*, 55(1), 123–146.
- 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.
- Zhao, H. (2014). Investigating teacher-supported peer assessment for EFL writing. *ELT Journal*, 68(2), 155–168.
- Zhao, H. (2018). New insights into the process of peer review for EFL writing: A process-oriented socio-cultural perspective. *Learning and Instruction*, 58, 263–273.
- Zheng, Y., & Cheng, L. (2008). College English Test (CET) in China. *Language Testing*, 25(3), 408–417.
- Zheng, Y., & Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL writing : A case study of Chinese lower-proficiency students. *Assessing Writing*, 37, 13–24.
- Zhou, J., Zheng, Y., & Tai, J. H.-M. (2020). Grudges and gratitude: The social-affective impacts of peer assessment. *Assessment & Evaluation in Higher Education*, 45(3),

345–358.

- Zhou, X., & Zhan, J. (2016). 从《要求》到《指南》：解读《大学英语教学指南》中的课程设置 [Understanding the changing College English curriculum in China: From the Requirements to the Guide]. *Foreign Languages in China*, 13(1), 13–18.
- Zhou, Y. (2013). 大学英语写作反馈方式的对比研究 [A contrastive research on feedback in teaching college English writing]. *Journal of the Foreign Language World*, (3), 87–96.
- Zhu, Q. (2018). *Chinese undergraduates' engagement with peer feedback: Perceptions and practices*. Unpublished PhD thesis, University of Hong Kong, HKSAR, China.
- Zhu, Q., & Carless, D. (2018). Dialogue within peer feedback processes: clarification and negotiation of meaning. *Higher Education Research and Development*, 37(4), 883–897.
- 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.

**Appendix A: Scoring Rubric Adapted from Jacobs et al.  
(1981)'s ESL Composition Profile**

Category, Score range & Criteria		Score
<b>Content</b>	<b>Excellent to very good (30-27):</b> knowledgeable, substantive, thorough development of thesis, relevant to assigned topic	
	<b>Good to average (26-22):</b> some knowledge of subject, adequate range, limited development of thesis, mostly relevant to the topic, but lacks detail	
	<b>Fair to poor (21-17):</b> limited knowledge of subject, little substance, inadequate development of topic	
	<b>Very poor (16-13):</b> does not show knowledge of subject, non-substantive, not pertinent, OR not enough to evaluate	
<b>Organization</b>	<b>Excellent to very good (20-18):</b> fluent expression, ideas clearly stated, well-organized, logical sequencing, cohesive	
	<b>Good to average (17-14):</b> somewhat choppy, loosely organized but main ideas stand out, logical but incomplete sequencing	
	<b>Fair to poor (13-10):</b> non-fluent, ideas confused or disconnected, lacks logical sequencing or development	
	<b>Very poor (9-7):</b> does not communicate, no organization, Or not enough to evaluate	
<b>Vocabulary</b>	<b>Excellent to very good (20-18):</b> sophisticated range, effective word/idiom choice and usage, word form mastery, appropriate register	
	<b>Good to average (17-14):</b> adequate range, occasional errors of word/idiom choice, usage, <i>but meaning not obscured</i>	
	<b>Fair to poor (13-10):</b> limited range, frequent errors of word/idiom choice and usage, <i>meaning confused or obscured</i>	
	<b>Very poor (9-7):</b> essentially translation, little knowledge of English vocabulary, idiom, word form, OR not enough to evaluate	

<b>Grammar</b>	<b>Excellent to very good (25-22):</b> effective complex constructions, few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions	
	<b>Good to average (21-18):</b> effective but simple constructions, minor problems in complex constructions, several errors of agreement, tense, number, word order/function, articles, pronouns, and prepositions, <i>but meaning seldom obscured</i>	
	<b>Fair to poor (17-11):</b> major problems in simple/complex constructions, frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, and deletions, <i>meaning confused or obscured</i>	
	<b>Very poor (10-5):</b> virtually no mastery of sentence construction rules, dominated by errors, does not communicate, OR not enough to evaluate	
<b>Mechanics</b>	<b>Excellent to very good (5):</b> demonstrates mastery of conventions, few errors of spelling, punctuation, capitalization, indentation/extra-spacing for paragraphing	
	<b>Good to average (4):</b> occasional errors of spelling, punctuation, capitalization, indentation/extra-spacing for paragraphing, <i>but meaning not obscured</i>	
	<b>Fair to poor (3):</b> Frequent errors of spelling, punctuation, capitalization, indentation/extra-spacing for paragraphing, <i>meaning confused or obscured</i>	
	<b>Very poor (2):</b> no mastery of conventions, dominated by errors of spelling, punctuation, capitalization, indentation/extra-spacing for paragraphing, OR not enough to evaluate	
<b>TOTAL</b>		

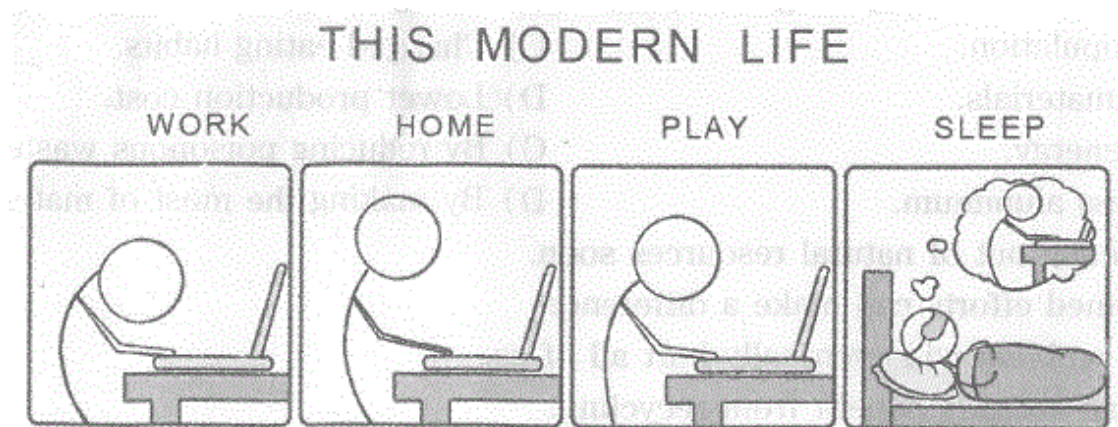
## Appendix B: Writing Prompts for the Six Writing tasks during the Intervention Period

### (Writing Task 1)

Directions: Write an essay commenting on the saying “Never go out there to see what happens, go out there to make something happen.” You can cite examples to illustrate the importance of being creative rather than mere onlookers in life. You should write at least 200 words.

### (Writing Task 2)

Directions: Write an essay based on the picture below. You should start your essay with a brief description of the picture and then comment on this kind of modern life. You should write at least 200 words.



### (Writing Task 3)

Directions: Write an essay based on the picture below. You should start your essay with a brief account of **the impact of the Internet on way people communicate** and then **explain whether electronic communication can replace face-to-face contact**. You should write at least 200 words.



"Dear Andy — How are you? Your mother and I are fine. We both miss you and hope you are doing well. We look forward to seeing you again the next time your computer crashes and you come downstairs for something to eat. Love, Mom and Dad."

**(Writing Task 4)**

Directions: For this part, you are allowed 30 minutes to write a short essay on **how to best handle the relationship between teachers and students**. You should write at least 200 words.

**(Writing Task 5)**

Directions: Write a short essay about **a campus activity that benefited you most**. You should state the reasons and write at least 200 words.

**(Writing Task 6)**

Directions: For this part, you are allowed 30 minutes to write a short essay. You should start your essay with a brief description of the picture and then express your views on the importance of reading literature. You should write at least 200 words.



*"Just think of it as if you're reading a long text-message."*



## Appendix C: Peer Feedback Form for Writing Task 6

Draft written by \_\_\_\_\_ Draft reviewed by \_\_\_\_\_ Date \_\_\_\_\_

### General Guidelines

Before providing feedback for your partner, make sure you read the instruction on coherence. Then you are required to read and comment on your read partner's draft guided by the some questions. As a final step, you are required to **FILL IN THE FORM**. Meanwhile, you can also give in-text comments and/or end-of-the text comments.

When you provide feedback for your partner, please bear in mind the following two points: Firstly, keep your feedback specific and constructive: You should not only point out the problems where they occur, but also offer clear advice on how to fix the problem. Secondly, be honest and friendly with your partner. The purpose of peer feedback is to help each other improve their draft.

### Instruction on cohesion

Creating **cohesion** means 'tying' our words, phrases, sentences and paragraphs together, to create a text where the relationships between these elements is clear and logical to the reader, giving the text 'flow'. We create cohesion at all these levels (word, phrase, sentence and paragraph), in order to direct our readers' attention to the development of our argument. Following are some of the cohesive devices (as underlined in the sentences) you can make use of.

**Repetition:** use a particular word or phrase across different sentences or paragraphs.

Example: Where does the designer get ideas and inspiration for new fashions? The fashion designer gets ideas and inspiration everywhere! Through television the designer experiences all the wonders of the entertainment world. In films the designer is exposed to the influences of all the arts and lifestyles throughout the world. Museum exhibits, art shows, world happenings, expositions, the theater, music, dance, and world travel are all sources of design inspiration to the fashion designer.

**Synonyms:** use a word or phrase in a later sentence which has the same or similar meaning to a keyword in the first sentence

Example: Excesses of fat beyond the body’s usual requirements become the storage fat that builds up as adipose or fatty tissue in many parts of the body. This is the type of fat that forms layers between the muscles and the skin and over the internal organs and other vital parts. These cushions do have some protective value, but this storage fat also contributes to obesity.

**Pronouns:** use a pronoun to refer back to a phrase already used

Example: When scientific experiments do not work out as expected, they are often considered failures until some other scientist tries them again. Those that work out better the second time around are the ones that promise the most rewards

**Transition words/expressions:**

<b>Adding similar ideas:</b> “in addition...”, “another reason is...”, and “equally...”
<b>Contrasting ideas:</b> “in contrast...”, “on the other hand...”, and “on the contrary...”, and “although...”
<b>Giving examples or evidence:</b> “for example...”, “for instance...”, and “to illustrate this...”
<b>Explaining results:</b> “as a result...”, “for this reason...”, and “consequently...”
<b>Sequencing:</b> “firstly...”, “secondly...”, “next...”, and “to begin with...”
<b>Providing explanations:</b> “owing to...”, “because of...”, and “due to...”
<b>Drawing conclusions:</b> “in conclusion...” and “to sum up...”

**Introducing your opinion:** “in my opinion...”, “to my mind...”, and “it seems to me that ...”

**Important tips:** Don’t over-use transition words or phrases or use them inappropriately, because you may confuse or irritate your readers. Show variety in cohesive devices.

### **Guiding questions for provision of feedback**

1. What are some strengths in this composition? Please list **AT LEAST ONE** strength and illustrate with examples from the composition.

#### **2. Feedback on cohesion**

(1) List **AT LEAST TWO** transition devices used in this composition and comment on this use?

(2) Other questions to consider:

What cohesive devices are used in this composition? Are they varied? What other devices can be used in this composition?

Do the transition words/expressions in this composition help you understand the logic of the ideas? Comment on inappropriately-used transition words/expressions and make suggestions.

#### **3. Feedback on other aspects**

(You are not required to provide feedback on every aspects listed below, but you are required to generate a total of **AT LEAST EIGHT** pieces of feedback)

- (1) Introductory paragraph: Is there a thesis statement in the introductory paragraph? Is the thesis statement congruent with the directions and well-developed through body paragraphs? Is the introductory paragraph appropriate in length?
- (2) Body paragraphs: Is there a topic sentence within each body paragraph? Is the topic sentence in each body paragraph well-supported? Is any part within the paragraph not related to the topic sentence, or any part within the paragraph not logically ordered? Is each body paragraph appropriate in length?
- (3) Concluding paragraph: Is there a conclusion in the composition? Does it echo to the thesis statement? Is the concluding paragraph appropriate in length?
- (4) Clarity in content: Is any part of the composition confusing?
- (5) Language use: What are some language problems in this composition?
- (6) What are other problems in this composition?

## 同伴反馈意见表（写作任务六）

### Peer Feedback Form for Writing Task 6 (Chinese)

作者\_\_\_\_\_ 评阅者\_\_\_\_\_ 评阅日期\_\_\_\_\_

#### 反馈流程及注意事项

请同学们按照以下三个步骤完成同伴反馈意见表：（1）认真学习反馈重点；（2）依据意见表所列的问题评阅同伴的作文；（3）填写意见表。在完成意见表的同时，也可以在文中或者文末做出标示。

请同学们在提供反馈意见时，特别注意以下两点：（1）反馈意见应具体：请指出问题并提供修改意见；（2）表达意见时，请注意语言表述，真诚友好。

#### 反馈重点

本轮写作任务的反馈重点是**衔接与连贯**。衔接是是实现语句、语篇连贯的重要手段。语句或者篇章如果缺乏必要的衔接，文章会显得杂乱无章、脉络不清。连贯是可以使语言与内容有机结合起来，使之符合逻辑，表达流畅。常见的衔接首段主要有以下四种（在举例中以下划线表示）。

**重复关键词：**在不同的句中重复关键词，比如，

Where does the designer get ideas and inspiration for new fashions? The fashion designer gets ideas and inspiration everywhere! Through television the designer experiences all the wonders of the entertainment world. In films the designer is exposed to the influences of all the arts and lifestyles throughout the world. Museum exhibits, art shows, world happenings, expositions, the theater, music, dance, and world travel are all sources of design inspiration to the fashion designer.

**同义词替换：**在后文中使用句中关键词的同义词，比如，

Excesses of fat beyond the body's usual requirements become the storage fat that builds up as adipose (脂肪的) or fatty tissue in many parts of the body. This is the type of fat that forms layers between the muscles and the skin and over the internal organs and

other vital parts. These cushions do have some protective value, but this storage fat also contributes to obesity (肥胖症).

**使用指代词：** 在后文中使用代词指代前文中提到的内容，比如：

When scientific experiments do not work out as expected, they are often considered failures until some other scientist tries them again. Those that work out better the second time around are the ones that promise the most rewards

**使用连接词：**

- 附加：“in addition...”, “another reason is...”, and “equally...”
- 比较：“in contrast...”, “on the other hand...”, and “on the contrary...”, and “although...”
- 举例：“for example...”, “for instance...”, and “to illustrate this...”
- 解释结果：“as a result...”, “for this reason...”, and “consequently...”
- 顺序：“firstly...”, “secondly...”, “next...”, and “to begin with...”
- 解释原因：“owing to...”, “because of...”, and “due to...”
- 结尾：“in conclusion...” and “to sum up...”
- 引出观点：“in my opinion...”, “to my mind...”, and “it seems to me that ...”

**特别注意**

- 衔接手段的使用应恰当适度，切忌盲目对堆砌连接词。
- 衔接手段的使用应多样化。

**反馈意见**

1. 这篇文章的优点有哪些？请**至少列举一处**并评价。
2. **衔接手段方面：**请列举这篇文章使用的至少**两处**衔接手段，并分析其使用是否恰当。

其他有关连贯和衔接方面的问题：

- (1) 这篇文章使用了哪些衔接手段？文章中衔接手段的应用多样吗？还有其他衔接手段可以应用吗？请举例说明
- (2) 文章中的连接词使用恰当吗，对于读者理解文章内容有帮助吗？请指出使用不恰当的连接词并提供修改意见。

**3. 其他方面**（请选择地评价以下几方面内容，但反馈意见的总数**不得低于八条**）：

- (1) 文章首段：文章的首段是否包含文章主题句？文章的主题句是否符合题目的要求？文章主题段落的内容是否围绕文章主题展开？首段的长处是否合适？
- (2) 主体段落：每个主体段落是否包含段落主题句？各段落中是否有偏离段落主题句的情况？各段落的句子安排是否符合逻辑？各段落的长度是否合适？
- (3) 结尾段落：文章是否有结尾段？结尾段落的内容和文章主题是否一致？结尾段落的长度是否合适？
- (4) 内容表达：文章的内容表达是否清楚？
- (5) 语言表达：文章种语言表达的问题有哪些？
- (6) 文章中的其他问题有哪些？

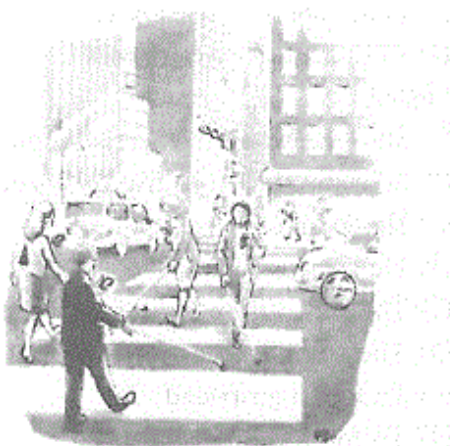
## Appendix D: Writing Prompts for Three Independent Writing Tasks during the Post-intervention Period

### (Writing Task 7)

Directions: Write an essay commenting on the saying ‘Learning is a daily experience and a lifetime mission.’ You can cite examples to illustrate the importance of lifelong learning. You should write at least 200 words.

### (Writing Task 8)

Directions: Write a short essay based on the picture below. You should start your essay with a brief account of the increasing use of the mobile phone in people's life and then explain the consequences of overusing it. You should write at least 200 words.



People are crossing the street looking at their cell phones and using walking sticks in order to see.

### (Writing Task 9)

Directions: Write a short essay about a course that has impressed you most in college. You should state the reasons and write at least 200 words.



## **Appendix E: Writing Prompts for and Pre-, Post-, and Delayed post-tests**

### **(Pre-test)**

Directions: Suppose you have two options upon graduation: one is to take a job in a company and the other to go to a graduate school. You are to make a choice between the two. Write an essay to explain the reasons for your choice. You should write at least 200 words within the 45 minutes.

### **(Post-test)**

Directions: Suppose you have two options upon graduation: one is to find a job somewhere and the other to start a business of your own. You are to make a decision. Write an essay to explain the reasons for your decision. You should write at least 200 words within the 45 minutes.

### **(Delayed post-test)**

Directions: Suppose you have two options upon graduation: one is to work in a state-owned business and the other in a joint venture. You are to make a choice between the two. Write an essay to explain the reasons for your choice. You should write at least 200 words within the 45 minutes.

## Appendix F: Error Coding Scheme

1. Do not distinguish among error types such as grammatical errors, lexical choice errors and mechanic errors.
2. Do not count identical errors within a text repeatedly.  
E.g.: First, go to a graduate school can show our ability...Second, go to a graduate school will enrich our knowledge. (“going to a graduate school”, one error)
3. Count errors separately, even though they concern a single word.  
E.g.: And they have made contributed to what they major in. (“contributions”, two errors, a word form error and a word choice error)
4. Count an expression that contains redundant vocabulary or lacks necessary vocabulary as an error.  
E.g.: In my opinion, a joint venture will give me more creative and challengeable tasks, so that I will be able to get more developing opportunities and benefit a lot from those tasks. (“developing” being redundant)
5. Count a wrongly-derived word as an error.  
E.g.: Secondly, going to a graduate school guarantees a higher-paying job. (“well-paid”)
6. Oral/informal expressions, including words, abbreviations, and idioms (e.g., cause and wanna) were counted as errors.
7. Count a misspelling as an error.  
E.g.: However, I still want to go to a graduate school to accept further education. (“further”)
8. Do not count errors in capitalization.  
E.g: The most impressive activity was Military Training. (“military training”, no error)
9. Accept both American and British spellings (e.g., center and centre; practice and practise).
10. Don not count an error that is incurred by another error.  
E.g.: Absolutely, countless knowledges attract me. (“Absolutely, countless knowledge attracts me”, one error)

11. Do not count punctuation mistakes, except for those that result in fragments or run-ons, wrong form of the possessive case, or those related to restrictive/non-restrictive relative clauses.

E.g.: If I start a business, most of my time will be filled with work, that's what I don't expect. (use of comma resulting in run-on sentence)

Starting a business means to take on a lot of responsibility which is a little difficult for me. (lack of comma required by non-restrictive relative clauses)

## Appendix G: Coding of revision episodes & cognitive operations in PF triggered revision episodes

Think-aloud Protocols	Revision episodes & cognitive operations
<p><b>(PF-4)</b> (1) My peer told me “<b>on the one hand</b>” and “<b>on the other hand</b>” were used to introduce two opposite statements. (2) I used “<b>on the one hand</b>” to introduce the idea of developing thinking ability, and “<b>on the other hand</b>” to introduce the idea of developing one’s perseverance. These two ideas were not opposite, so the two linking phrases were inappropriate. (3) What other phrases can I use, “<b>for one thing</b>” and “<b>for another thing</b>”? Do they introduce opposite statements? (4) [Reading a sentence in her first draft] “<b>Literature helps to develop the comprehensive and critical thinking in one’s mind.</b>” (5) [Making the decision to delete “on one hand” in the sentence “On one hand, by providing systematically knowledge in some fields, literature helps to develop the comprehensive and critical thinking in one’s mind.”] I can start the sentence with “<b>by</b>” directly and delete “<b>on the one hand</b>”. (6) There is no need for the linking device. This sentence clearly introduces one argument. (7) Since “<b>on the one hand</b>” is deleted, I need to replace “<b>on the other hand</b>” with another phrase. (8) OK, I use “<b>besides</b>”. (9) I wrote it down. //</p> <p><b>(PF-5)</b> (1) She [His peer] crossed out “<b>ly</b>” [in the word “systematically”], (2) I will use “<b>systematical</b>”. //</p> <p><b>(Other)</b> [Reading] “<b>By providing systematical knowledge in some fields, literature helps to develop comprehensive and critical thinking in one’s mind</b>”, this sentence is the topic sentence for the second paragraph. Is “<b>develop comprehensive and critical thinking in one’s mind</b>” consistent with “<b>provide</b></p>	<p><b>PF-triggered revision episode 4</b></p> <ol style="list-style-type: none"> <li>(1) Referring to PF</li> <li>(2) Justifying</li> <li>(3) Retrieving L2 knowledge</li> <li>(4) Referring to text</li> <li>(5) Generating changes</li> <li>(6) Evaluating</li> <li>(7) Monitoring</li> <li>(8) Retrieving L2 knowledge</li> <li>(9) Generating changes</li> </ol> <p><b>PF-triggered revision episode 5</b></p> <ol style="list-style-type: none"> <li>(1) Referring to PF</li> <li>(2) Generating changes</li> </ol> <p><b>Other revision episode</b> (Not included in the data &amp; not analyzed for cognitive operations)</p>

**knowledge completely**” [the revised thesis statement]? **“develop comprehensive and critical thinking in one’s mind”** was related to what I originally wrote, **“train a person how to think”** [in his first draft]. Also, I did not include any ideas about **“critical”** in this paragraph. I deleted the **“critical”** part. //

**(PF-6)** (1) Here is a misspelling [the “write”]. My peer added “r”. (2) Yes, the “writer”. I corrected it. //

**(PF-7)** (1) [Reading] **“For the most aspects of the topic the writer wants to discuss”**, my peer asked to move **“in a book”** [in the next sentence] to the beginning of this sentence. (2) I added **“in a book”** [as his peer suggested]. (3) The sentence is very coherent. //

**(Other)** [Reading] **“Enough space is provided”**, I have used **“provided”** [in the first paragraph]. I will use another word, **“given”**. **“Given to”** or **“given for”**? **“Given for”** does not sound right. I will write **“Enough space is given to most aspects of the topic the writer wants to discuss”**. //

**(Other)** [Reading] **“So, the concepts can be described detailedly”**, this sentence seems to be incomplete. In a book, there are concepts to be explained. What other things are there in the book? The concepts and evidence. Yes, in English, **“evidence”**. **So, the concepts can be described detailedly and evidence 随处可见**[meaning is everywhere]. I remembered for a previous writing task, my peer used **“within easy reach”**, I can use it here, **“evidence is within easy reach”**. If I still use **“while”** in this sentence, the sentence will be long. I can use two sentences. I start a new sentence with **“however”** [**So, the concepts can be described detailedly and evidence is within easy reach. However, in a short message, everything must be written briefly**]. //

**(Other)** **“In a short message”**, I should not use it [“short message”], because it means text message, about mobile phones. I can use **“articles”**, **“articles in the internet.”** //

**PF-triggered revision episode 6**

- (1) Referring to PF
- (2) Generating changes

**PF-triggered revision episode 7**

- (1) Referring to PF
- (2) Generating changes
- (3) Evaluating

**Other revision episode**

(Not included in the data & not analyzed for cognitive operations)

**Other revision episode**

(Not included in the data & not analyzed for cognitive operations)

**Other revision episode**

(Not included in the data & not analyzed for cognitive operations)

<p><b>(PF-8)</b> (1) My peer added <b>“overall”</b> [“to reach an overall” conclusion]. (2) I already compared it with <b>“complete”</b> [when he was revising the first paragraph]. <b>“Overall”</b> can be used as a noun and an adjective; <b>“complete”</b> means “including all parts, with nothing missing”. (3) I will use <b>“complete”</b> here. Its meaning is more appropriate than <b>“overall”</b>. //</p> <p><b>(Other)</b> [Reading] <b>“At the same time, it’s hard to draw a complete conclusion with limited knowledge in the message, but the reader will think critically and reach a complete conclusion after he reads literature since he gains everything he needs”</b>, I used <b>“everything”</b> previously [in “everything must be written briefly” in his first draft]. So here I revised it into “he gains what he needs”. //</p>	<p><b>PF-triggered revision episode 8</b></p> <ul style="list-style-type: none"> <li>(1) Referring to PF</li> <li>(2) Retrieving L2 knowledge</li> <li>(3) Generating changes</li> <li>(4) Evaluating</li> </ul> <p><b>Other revision episode</b> (Not included in the data &amp; not analyzed for cognitive operations)</p>
---	--

## Appendix H: The Perceptions of Peer Feedback Questionnaire

### 同伴互评认知态度问卷

本问卷一共 9 个开放性问题。对于每一个问题，请从 1-5 中勾选一个分数以表示自己的同意程度，并对自己的选择进行详细解释。你的答案没有对错之分，请根据目前的情况真实并详细作答，谢谢配合！

This questionnaire contains nine open-ended questions. Please first circle corresponding number from 1 to 5 on the Likert Scale to show how much you agree or disagree with the following statements and then explain for your choice. This is not a test so there are no right or wrong answers. We are only interested in your personal opinion. The result of this questionnaire will be used for research purpose so your sincere answers to these questions ensure the success of this research project. Thank you!

代 码 (Code) : \_\_\_\_\_ 专 业 (Major) : \_\_\_\_\_

性 别 (Gender) :  男 (Male)  女 (Female)

1. 通过参与一学期的同伴互评活动，我在作文的内容方面有所提高。  
By participating in peer feedback activities all through the semester, I have improved in the content of my composition.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

2. 通过参与一学期的同伴互评活动，我在作文的结构方面有所提高。

By participating in peer feedback activities all through the semester, I have improved in the organization of my composition.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

3. 通过参与一学期的同伴互评活动，我在作文的语言使用方面有所提高。

By participating in peer feedback activities all through the semester, I have improved in the language use (including grammar, vocabulary and sentence structure) of my composition.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

4. 学期初，老师对于同伴互评的培训，我从中受益。

I found the general training session on peer feedback implemented at the beginning of this semester useful.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>



请具体阐述 (Please elaborate on your choice)

5. 在每次写作任务中，老师对于同伴互评的针对性培训，我从中受益。  
I found the six focused training sessions on peer feedback implemented within each writing task useful.

1

2

3

4

5

非常不同意

不同意

中立

同意

非常同意

**Strongly disagree**

**Disagree**

**Neither disagree or agree**

**Agree**

**Strongly agree**

请具体阐述 (Please elaborate on your choice)

6. 每次写作任务中，评价同伴作文这一环节，我从中受益。  
I benefitted from reviewing my peers' compositions throughout the semester.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

7. 每次写作任务中，同伴为我的作文提供修改意见，我从中受益。  
I benefitted from receiving feedback from my peers throughout the semester.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

8. 每次写作任务中，讨论同伴意见这一环节，我从中受益。  
I found the paired discussions on feedback within each writing task was helpful.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

9. 在每次写作任务中使用同伴反馈意见表，我从中受益。  
I found using the PF forms throughout the semester useful.

1	2	3	4	5
非常不同意	不同意	中立	同意	非常同意
<b>Strongly disagree</b>	<b>Disagree</b>	<b>Neither disagree or agree</b>	<b>Agree</b>	<b>Strongly agree</b>

请具体阐述 (Please elaborate on your choice)

## **Appendix I: Questions for the Semi-structured Interviews**

1. Throughout this semester, you have participated in peer feedback activities. What is your opinion about peer feedback? How is it different from teacher feedback?
2. How do you feel about the change in your writing performance in expository writing? What factors can account for your change or no change in your writing performance? Of the three aspects in writing performance, content, organization, and language use, which aspect do you think is improved most obviously?
3. How do you like the general training implemented at the beginning of the semester? Do you find it helpful and necessary?
4. How do you like the focused training sessions embedded in each writing cycle?
5. Do you benefit from reading your peers' compositions and giving feedback to your peers? If so, what are the benefits? If not, why not?
6. Do you benefit from receiving feedback from your peers? If so, what are the benefits? If not, why not?
7. Do you feel it necessary to include a discussion about feedback within each writing task? Will written peer feedback suffice?
8. Do you benefit from using PF forms in peer feedback activities? If so, what are the benefits? If not, why not?
9. If you have the choice, would you prefer to work with the same partner for a whole semester or to work with different partners?
10. What types of feedback do you prefer in giving feedback to others?
11. What types of feedback do you expect from others?
12. Will you be able to recall and apply skills acquired in peer feedback activities to the following compositions?
13. Did you manage to fulfil the writing tasks and peer feedback activities?
14. Would you like peer feedback to be implemented in College English course in the following semesters?

# Appendix J: Ethics Forms

## PARTICIPANT INFORMATION SHEET

(Faculty Dean)

**Research project:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

### Research Introduction

I am Zhiqing Xu, a PhD candidate in the Faculty of Education and Social Work, the University of Auckland, New Zealand. I am conducting research on the effects of peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' text revisions and writing performance.

Peer feedback is a collaborative activity in which students engage in reading, critiquing and providing feedback on each other's writing. In teaching writing to non-English majors through College English course, peer feedback constitutes a valuable component, with the potential of helping students improve their writing.

This project will take place in the spring semester 2018 and summer vacation of 2018, from March 2017 to September 2018, consisting of two parts: a quasi-experimental study and a multiple case study taking place simultaneously.

### Faculty Involvement

Your permission and cooperation are the prerequisites of conducting this project. I am seeking your permission to get access to EFL teachers and students. Once granted permission, I will contact the secretary of your faculty to approach teachers and students to invite them to participate in my research on my behalf and pass on the attached research information sheets and consent forms. In total, this research plans to recruit one teacher participant and 50 student participants from each of the two classes the teacher participant is teaching (including six for the multiple case study from the intervention group).

In addition, I would like to request for your assurance that both the students' participation and the teacher's participation in my research is voluntary. For the student participants, their non-participation or withdrawal will have no effect on their grades or relationships with the faculty; for the teacher participant, his/her non-participation or withdrawal will have no effect on his/her career or relationship with the faculty.

### **Teacher Involvement**

At the beginning of the spring semester 2018, the researcher and the teacher participant will decide which class will use peer feedback activities and which class will use regular teacher feedback activities.

In the peer feedback class, the teacher will implement a general peer feedback training session, and six writing tasks with peer feedback activities throughout the semester. Peer feedback training will be implemented through a 90-minute general training session and six focused training sessions embedded in the six writing tasks. Each writing task will last two weeks. Within each writing task, in the first week, students write their first drafts, exchange their drafts in pairs and provide written feedback on each other's first draft after class; in the second week, they participate in a focused peer feedback training session and a paired discussion session in class, and finally write a second draft after class. To facilitate the intervention, the teacher participant will be provided with peer feedback training, and be reminded to implement the intervention in one class, and keep his/her regular teacher feedback activities for the six writing tasks in the other class.

In addition, the teacher participant will assist the researcher in administering tests and questionnaires. In Week 1, 16, and 25, the teacher will administer the pre-test, post-test, and delayed post-test respectively, to all the student participants in the two classes. In Week 16, the teacher will also administer the Perceptions of Peer feedback Questionnaire to student participants in the peer feedback class. In Week 25, the delayed post-test will be administered to all student participants. Each test will take 45 minutes and the questionnaire will take 30 minutes.

### **Student Involvement**

Student participants recruited from the two classes will participate in the pre-test, post-test, and delayed post-test in Week 1, 16, and 25, respectively. In addition, only student participants in the peer feedback class will answer the Perceptions of Peer feedback Questionnaire in Week 16. Finally, the six case study participants recruited from the peer feedback class will participate in

some additional activities, including a think-aloud training session, two think-aloud sessions in Writing Task 5 and 6, and individual semi-structured interviews after the post-test. Both the think-alouds and semi-structured interviews will be conducted in Chinese.

As regards time commitment from student participants, the total time required for each participant in the peer feedback class will be 2 hours and 45 minute; the total time for each participant in the teacher feedback class will be 2 hours and 15 minutes; and the total time for each of the six case study participant will be 6 hour and 55 minutes. In order to ensure participants in the teacher feedback group are not disadvantaged, the teacher participant will provide them with extensive peer feedback training and all resources used by the peer feedback class after the study.

### **Participants' Rights to Withdraw**

Participants will be informed of the purpose and procedure of the current research through the Participant Information Sheet and the Consent Form, and they are completely voluntary and entitled to withdraw either themselves or any data provided by them without giving any reasons within up to three weeks after the data collection is completed. After that time any withdrawal of data could not be possible as analysis would be underway.

### **Confidentiality**

The confidentiality of participants will be guaranteed during the whole research process. For the quasi-experimental study, each student participant will be given an identification code to use throughout the research. The purpose of using identification codes is to link questionnaires to the written texts in the pre-test, post-test, and delayed post-test. As for participants in the multiple case study, each of them will be provided a pseudonym. All the coding information will be kept separately from the data, and will only be known to the researcher. Participants' names and other facts that might identify them will not appear when we present this study or publish its results. Any identifying information about the university and the faculty will be removed. No identifiable information will be released to a third party.

### **Data Storage, Retention, Destruction and Future Use**

The collected Consent Forms and hard copy data will be securely stored in a locked cabinet at the University of Auckland, and electronic data will be stored confidentially in the researcher's computer, which is password-protected. After six years, all hard copy data will be shredded and the digital information will be deleted permanently from all electronic devices. The collected

data will be primarily presented in the researcher's PhD thesis and may also be used for future academic publications or conference presentations.

Thank you for sharing your time to read this information sheet. If you have any inquiries or questions, please feel free to contact anyone in the following contact list.

### Contact details

Researcher	Main supervisor	Co-supervisor
<p>Zhiqing Xu</p> <p>School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:zxu621@aucklanduni.ac.nz">zxu621@aucklanduni.ac.nz</a></p> <p>Ph: +86 431 85929518</p>	<p>Professor Lawrence Jun Zhang</p> <p>School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:lj.zhang@auckland.ac.nz">lj.zhang@auckland.ac.nz</a></p> <p>Ph: +64 9 373 7999 ext 48750</p>	<p>Professor Judy Margaret Parr</p> <p>School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:jm.parr@auckland.ac.nz">jm.parr@auckland.ac.nz</a></p> <p>Ph: +64 9 923 8998</p>

You may also contact the head of the School of Curriculum and Pedagogy, Associate Professor Helen Hedges at [h.hedges@auckland.ac.nz](mailto:h.hedges@auckland.ac.nz) or +64 9 373 7999 ext 48606.

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142, New Zealand. Telephone 09 373-7999 ext. 83711. Email: [ro-ethics@auckland.ac.nz](mailto:ro-ethics@auckland.ac.nz).

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.



## **PARTICIPANT INFORMATION SHEET**

### **(Student participant in the quasi-experimental study)**

**Research project:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

#### **Research Introduction**

I am Zhiqing Xu, a PhD candidate in the Faculty of Education and Social Work, the University of Auckland, New Zealand. I am conducting research on the effects of peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' text revisions and writing performance.

Peer feedback is a collaborative activity in which students engage in reading, critiquing and providing feedback on each other's writing. In teaching writing to non-English majors through College English course, peer feedback constitutes a valuable component, with the potential of helping students improve their writing.

This project will take place in the spring semester 2018 and summer vacation of 2018, from March 2017 to September 2018, consisting of two parts: a quasi-experimental study and a multiple case study taking place simultaneously.

I would like to invite you to participate in this research. I have contacted your faculty and gained permission from the dean to ask for your involvement.

#### **Student Involvement**

If you consent to participate in the quasi-experimental study, you, together with other classmates in your class, will be assigned to one of two conditions (peer feedback class and teacher feedback class). You will be invited to participate in the follow activities.

In Week 1, 16, and 25, you will take part in the pre-test, post-test, and delayed post-test respectively. Each tests involves a given-topic writing tasks, taking you about 45 minutes to complete. Your compositions in the three tests will be collected. In addition, your two drafts for Writing Task 6 will also be collected.

If you are in the peer feedback class, you will also need to fill in the Perceptions of Peer Feedback Questionnaire in Week 16. It will take you approximately 30 minutes to complete the questionnaire.

### **Compensation**

In order to express the researcher's gratitude for your participation, upon completion of the study, you will receive a 50 RMB (NZ \$10) shopping coupon as a token of appreciation.

When the research finishes, a summary of the findings will also be made available to you. Please leave your email address in the Consent Form to request a summary of research findings.

### **Participants' Rights to Withdraw**

Participation in this research is entirely voluntary. The faculty Dean has given assurance that your participation or non-participation will not affect your grades or your relationship with the faculty.

You are entitled to withdraw either yourselves or any data you have provided without giving any reasons within up to three weeks after the data collection is completed. After that time any withdrawal of data could not be possible as data analysis would be underway. The Dean has given assurance that your withdrawal will not affect your grades or your relationship with the faculty.

### **Confidentiality**

Your identity as well as any information of you that is shared with the researcher will remain confidential. You will be given an identification code for use in questionnaires and in writing tests. The purpose of using the identification code is to link the data you have provided, and the information is kept separately from the data, known to the researcher only. Your anonymity may not be assured during the research, but the researcher will keep your identity confidential during the research process and that no identifying information will be included in any research outputs. No identifiable information will be released to a third party.

### **Data Storage, Retention, Destruction and Future Use**

The collected Consent Forms and hard copy data will be securely stored in a locked cabinet at the University of Auckland, and electronic data will be stored confidentially in the researcher's computer, which is password-protected. After six years, all hard copy data will be shredded and

the digital information will be deleted permanently from all electronic devices. The collected data will be primarily presented in the researcher's PhD thesis and may also be used for future academic publications or conference presentations.

Thank you for sharing your time to read this information sheet. If you have any inquiries or questions, please feel free to contact anyone in the following contact list.

### Contact details

Researcher	Main supervisor	Co-supervisor
<p>Zhiqing Xu</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:zxu621@aucklanduni.ac.nz">zxu621@aucklanduni.ac.nz</a></p> <p>Ph: +86 431 85929518</p>	<p>Professor Lawrence Jun Zhang</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:lj.zhang@auckland.ac.nz">lj.zhang@auckland.ac.nz</a></p> <p>Ph: +64 9 373 7999</p> <p>ext 48750</p>	<p>Professor Judy Margaret Parr</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:jm.parr@auckland.ac.nz">jm.parr@auckland.ac.nz</a></p> <p>Ph: +64 9 923 8998</p>

You may also contact the head of the School of Curriculum and Pedagogy, Associate Professor Helen Hedges at [h.hedges@auckland.ac.nz](mailto:h.hedges@auckland.ac.nz) or +64 9 373 7999 ext 48606.

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142, New Zealand. Telephone 09 373-7999 ext. 83711. Email: [ro-ethics@auckland.ac.nz](mailto:ro-ethics@auckland.ac.nz).

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

# **PARTICIPANT INFORMATION SHEET**

## **(Case study participant)**

**Research project:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

### **Research Introduction**

I am Zhiqing Xu, a PhD candidate in the Faculty of Education and Social Work, the University of Auckland, New Zealand. I am conducting research on the effects of peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' text revisions and writing performance.

Peer feedback is a collaborative activity in which students engage in reading, critiquing and providing feedback on each other's writing. In teaching writing to non-English majors through College English course, peer feedback constitutes a valuable component, with the potential of helping students improve their writing.

This project will take place in the spring semester 2018 and summer vacation of 2018, from March 2017 to September 2018, consisting of two parts: a quasi-experimental study and a multiple-case study taking place simultaneously.

I would like to invite you to participate in this research. I have contacted your faculty and gained permission from the dean to ask for your involvement.

### **Student Involvement**

If you consent to participate, you will be invited to take part in the three tests and complete the Perceptions of Peer feedback Questionnaire, as well as participate in the following activities: a think-aloud training session (approximately 60 minutes), two think-aloud sessions in Writing Task 5 and 6 (each approximately 60 minutes), and a semi-structured interview (approximately 30 minutes).

In addition, the following documents will be collected from you as a participant:

- (1) Your first draft, PF form, and second draft for Writing Task 6 will be collected, photocopied, and returned to you at the end of the study.
- (2) The paired discussion you are in for Writing Task 6 will be audio-recorded and transcribed by the researcher. You will be provided with a copy of the transcripts for review, editing, deletion, and final approval. They will then be translated into English by the researcher, and finally analyzed based on research questions.
- (3) Your think-aloud protocols for Writing Task 6 will be audio-recorded. You will be provided with a copy of the transcripts for review, editing, deletion, and final approval. They will then be translated into English by the researcher, and finally analyzed based on research questions.
- (4) The semi-structured interview you participate in will be audio-recorded. The researcher will conduct, audio-record and transcribe each interview individually. You will be provided with a copy of the transcription for review, editing, deletion, and final approval. The transcripts will then be translated into English by the researcher, and finally analyzed based on research questions.

### **Compensation**

When the research finishes, a summary of the findings will be made available to you. Please leave your email address in the Consent Form to request a summary of research findings.

In order to express the researcher's gratitude for your participation in this research, you will be provided with a souvenir valued at 50 RMB (NZ \$10) and a 50 RMB (NZ \$10) shopping coupon as a token of appreciation.

### **Participants' Rights to Withdraw**

Participation in this research is entirely voluntary. The faculty Dean has given assurance that your participation or non-participation will not affect your grades or your relationship with the faculty.

You are entitled to withdraw either yourselves or any data you provided without giving any reasons within up to three weeks after the data collection is completed. After that time any withdrawal of data could not be possible as data analysis would be underway. The faculty Dean has given assurance that your participation or non-participation will not affect your grades or your relationship with the faculty.

### **Confidentiality**

Your identity as well as any information of you that is shared with the researcher will remain confidential. You will be provided a pseudonym when participating in the research project, and we will use this pseudonym rather than your name in all related data to protect your identity. All the coding information is kept separately from the data, and will only be known to the researcher. No identifiable information will be released to a third party.

### **Data Storage, Retention, Destruction and Future Use**

The collected Consent Forms and hard copy data will be securely stored in a locked cabinet at the University of Auckland, and electronic data will be stored confidentially on the researcher's computer, which is password-protected. After six years, all hard copy data will be shredded and the digital information will be deleted permanently from all electronic devices. The collected data will be primarily presented in the researcher's PhD thesis and may also be used for future academic publications or conference presentations.

Thank you for sharing your time to read this information sheet. If you have any inquiries or questions, please feel free to contact anyone in the following contact list.

### **Contact details**

<b>Researcher</b>	<b>Main supervisor</b>	<b>Co-supervisor</b>
<p>Zhiqing Xu</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:zxu621@aucklanduni.ac.nz">zxu621@aucklanduni.ac.nz</a></p> <p>Ph: +86 431 85929518</p>	<p>Professor Lawrence Jun Zhang</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:lj.zhang@auckland.ac.nz">lj.zhang@auckland.ac.nz</a></p> <p>Ph: +64 9 373 7999</p> <p>ext 48750</p>	<p>Professor Judy Margaret Parr</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:jm.parr@auckland.ac.nz">jm.parr@auckland.ac.nz</a></p> <p>Ph: +64 9 923 8998</p>

You may also contact the head of the School of Curriculum and Pedagogy, Associate Professor Helen Hedges at [h.hedges@auckland.ac.nz](mailto:h.hedges@auckland.ac.nz) or +64 9 373 7999 ext 48606.

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142, New Zealand. Telephone 09 373-7999 ext. 83711. Email: [ro-ethics@auckland.ac.nz](mailto:ro-ethics@auckland.ac.nz).

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

## **PARTICIPANT INFORMATION SHEET**

### **(Teacher participant)**

**Research project:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

#### **Research Introduction**

I am Zhiqing Xu, a PhD candidate in the Faculty of Education and Social Work, the University of Auckland, New Zealand. I am conducting research on the effects of peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' text revisions and writing performance.

Peer feedback is a collaborative activity in which students engage in reading, critiquing and providing feedback on each other's writing. In teaching writing to non-English majors through College English course, peer feedback constitutes a valuable component, with the potential of helping students improve their writing.

This project will take place in the spring semester 2018 and summer vacation of 2018, from March 2017 to September 2018, consisting of two parts: a quasi-experimental study and a multiple-case study taking place simultaneously.

I would like to invite you to participate in this research. I have contacted your faculty and gained permission from the dean to ask for your involvement.

#### **Teacher Involvement**

If you consent to participate, I would like to incorporate six writing tasks with peer feedback activities in your College English course. For the two classes you teach, you will implement peer feedback activities in one class, while implement your regular teacher feedback in the other class.

In the peer feedback class, you will implement a general peer feedback training, and six writing tasks with peer feedback activities as designed by the researcher throughout the semester. Peer feedback training will be implemented through a 90-minute general training session and six focused training sessions embedded in the six writing tasks. Each writing task will last two weeks. Within each writing task, in the first week, students write their first drafts, exchange their



drafts in pairs and provide written feedback on each other's first draft after class; in the second week, they participate in a focused peer feedback training session and a paired discussion session in class, and finally write a second draft after class. To facilitate the intervention, I would also like you to join peer feedback training with me before the intervention, and have regular discussions with me during the intervention.

In the teacher feedback class, you will implement your regular teaching procedures and maintain your regular feedback practices. You will ask the students to write on the same six topics in two drafts and provide them with your feedback on their first drafts.

In addition, you will conduct the following activities:

- (1) In Week 1, 16, and 25, you will administer the pre-test, post-test, and delayed post-test respectively, to all the student participants in the two classes. In Week 16, you will also administer the Perceptions of Peer feedback Questionnaire to student participants in the peer feedback class.
- (2) In Week 14, you will collect the two drafts of Writing Task 6 from the student participants in both the peer feedback class and teacher feedback class.
- (3) In Week 19, 21, and 23, you will need to assign three writing tasks to the students participants in both the peer feedback class and teacher feedback class.

### **Compensation**

In order to express the researcher's gratitude for your participation, upon completion of the study, you will receive a 50 RMB (NZ \$10) shopping coupon as a token of appreciation.

When the research finishes, a summary of the findings will also be made available to you. Please leave your email address in the Consent Form to request a summary of research findings.

### **Participants' Rights to Withdraw**

Participation in this research is entirely voluntary. The faculty Dean has given assurance that your participation or non-participation will not affect your career or your relationship with the faculty.

You are entitled to withdraw either yourselves or any data you provided without giving any reasons within up to three weeks after the data collection is completed. After that time any withdrawal of data could not be possible as data analysis would be underway. The faculty Dean

has given assurance that your participation or non-participation will not affect your career or your relationship with the faculty.

### **Confidentiality**

Your identity as well as any information of you that is shared with the researcher will remain confidential. Your anonymity may not be assured during the research, but the researcher will keep your identity confidential during the research process and that no identifying information will be included in any research outputs. No identifiable information will be released to a third party.

### **Data Storage, Retention, Destruction and Future Use**

The collected Consent Forms and hard copy data will be securely stored in a locked cabinet at the University of Auckland, and electronic data will be stored confidentially on the researcher's computer, which is password-protected. After six years, all hard copy data will be shredded and the digital information will be deleted permanently from all electronic devices. The collected data will be primarily presented in the researcher's PhD thesis and may also be used for future academic publications or conference presentations.

Thank you for sharing your time to read this information sheet. If you have any inquiries or questions, please feel free to contact anyone in the following contact list.

### **Contact details**

<b>Researcher</b>	<b>Main supervisor</b>	<b>Co-supervisor</b>
<p>Zhiqing Xu</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:zxu621@aucklanduni.ac.nz">zxu621@aucklanduni.ac.nz</a></p> <p>Ph: +86 431 85929518</p>	<p>Professor Lawrence Jun Zhang</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:lj.zhang@auckland.ac.nz">lj.zhang@auckland.ac.nz</a></p> <p>Ph: +64 9 373 7999</p> <p>ext 48750</p>	<p>Professor Judy Margaret Parr</p> <p>School of Curriculum and Pedagogy,</p> <p>Faculty of Education and Social Work,</p> <p>The University of Auckland,</p> <p>Gate 3, 74 Epsom Ave, Auckland</p> <p><a href="mailto:jm.parr@auckland.ac.nz">jm.parr@auckland.ac.nz</a></p> <p>Ph: +64 9 923 8998</p>

You may also contact the head of the School of Curriculum and Pedagogy, Associate Professor Helen Hedges at [h.hedges@auckland.ac.nz](mailto:h.hedges@auckland.ac.nz) or +64 9 373 7999 ext 48606.

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142, New Zealand. Telephone 09 373-7999 ext. 83711. Email: [ro-ethics@auckland.ac.nz](mailto:ro-ethics@auckland.ac.nz).

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

## CONSENT FORM

(Faculty Dean)

**(THIS FORM WILL BE HELD FOR A PERIOD OF SIX YEARS)**

**Project title:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

I have read the Participant Information Sheet and have understood the nature of the research being undertaken by Zhiqing Xu. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree to provide research sites.
- I agree to allow the department secretary to help with this research.
- I agree to allow undergraduate students to join this research.
- I agree to allow teachers to join this research.
- I agree to allow the teacher participant to use peer feedback activities within the College English course for the spring semester 2018.
- I understand that participation in this research project is voluntary.
- I assure that participation, non-participation or withdrawal will not affect career, employment, grade, academic performance, and relationships with the faculty.
- I understand that participants will be asked to use a unique identification code on their questionnaires and all written texts in the pre-test, post-test, and delayed test. The coding information will only be known to the researcher. Participants' confidentiality will be guaranteed.
- I understand that if the information provided by the six case study participants is to be reported/published, pseudonyms will be used to protect their identity and confidentiality will be assured.
- I understand that hard copy and digital data will be stored separately and securely for a period of six years and then destroyed.
- I understand that any identifying details about the university and faculty will be removed.
- I understand that no identifying information will be disclosed to a third party or the public.

- I wish to receive a copy of the research findings by email \_\_\_\_\_ (If not, please leave this blank).

Name \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS  
COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

## CONSENT FORM

**(Student participant for the quasi-experimental study)**

**(THIS FORM WILL BE HELD FOR A PERIOD OF SIX YEARS)**

**Project title:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

I have read the Participant Information Sheet and have understood the nature of the research being undertaken by Zhiqing Xu. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree to participate in the research.
- I understand that participation in this research project is completely voluntary.
- I agree to be assigned to either the intervention group or the comparison group.
- I agree to take three writing tests and answer the Perceptions of Peer Feedback Questionnaire if I am in the intervention group; and to take three writing tests if I am in the comparison group.
- I agree to provide my two drafts for Writing Task 6 for research purposes.
- I understand that I will not be disadvantaged as participants in the comparison group. I will be provided with extensive peer feedback training and all resources used by the treatment group after the study.
- I understand that I will be asked to use a unique identification code, which can only be known to the researcher, in the questionnaire and the three writing tests.
- I understand that the faculty Dean has given the assurance that my participation, non-participation or withdrawal will have no effects on my grades or relationship with the faculty.
- I understand that I have the right to withdraw either myself or any data provided by me without giving any reasons within up to three weeks after the data collection is completed.
- I understand that hard copy and digital data will be stored separately and securely for a period of six years and then destroyed.

- I understand that the data collected from the research will be used for the researcher's PhD thesis at the University of Auckland, and may be used for academic publications, and conference presentations.
- I know my confidentiality will be protected by the researcher and I will not be identified in any of the researcher's publications.
- I wish to receive a copy of the research findings by email \_\_\_\_\_ (If not, please leave this blank).

Name \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

## CONSENT FORM

(Case study participant)

(THIS FORM WILL BE HELD FOR A PERIOD OF SIX YEARS)

**Project title:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

I have read the Participant Information Sheet and have understood the nature of the research being undertaken by Zhiqing Xu. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree to participate in the research.
- I understand that participation in this research project is completely voluntary.
- I agree to provide my first draft, the corresponding PF form, and my second draft for Writing Task 5 to the researcher for research purposes.
- I agree to participate in a think-aloud training session and two think-aloud sessions.
- I agree to participate in a semi-structured interview.
- I understand that the discussion session, the think-aloud session I participate in in Writing Task 6, and the semi-structured interview after the post-test will be audio-recorded.
- I understand that the faculty Dean has given the assurance that my participation, non-participation or withdrawal will have no effects on my grades or relationship with the faculty.
- I understand that I can refuse to answer any questions in the interview, and may request the recording to be stopped anytime without having to give any reasons.
- I understand that the researcher will conduct and transcribe the audio-recordings individually, and a copy of them will be provided to me individually for review, editing, deletion, and final approval. The transcripts will then be translated into English by the researcher, and finally analyzed based on research questions.
- I understand that I have the right to withdraw either myself or any data provided by me without giving any reasons within up to three weeks after the data collection is completed.



- I understand that if the information provided by me is to be reported/published, a pseudonym will be used to protect my identity and confidentiality will be assured.
- I understand that hard copy and digital data will be stored separately and securely for a period of six years and then destroyed.
- I understand that the data collected from the research will be used for the researcher's PhD thesis at the University of Auckland, and may be used for academic publications, and conference presentations.
- I know my confidentiality will be protected by the researcher and I will not be identified in any of the researcher's publications.
- I wish to receive a copy of the research findings by email \_\_\_\_\_ (If not, please leave this blank).

Name \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.

## **CONSENT FORM**

**(Teacher participant)**

**(THIS FORM WILL BE HELD FOR A PERIOD OF SIX YEARS)**

**Project title:** The Effects of Peer feedback on Chinese English-as-a-Foreign Language (EFL) Learners' Text Revisions and Writing Performance

I have read the Participant Information Sheet and have understood the nature of the research being undertaken by Zhiqing Xu. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree to participate in the research.
- I understand that participation in this research project is completely voluntary.
- I understand that the faculty Dean has given the assurance that my participation, non-participation or withdrawal will have no effects on my career or relationship with the faculty.
- I understand that I will be provided with training prior to the intervention and support during the intervention by the researcher.
- I understand that I will incorporate peer feedback in my writing instruction in one class. For the same period of time, I will keep my regular teacher feedback practices in the other class.
- I understand that I will not provide any form of feedback to any of the participants' compositions during the six-week post-intervention period.
- I agree to help collect data from student participants in my class.
- I understand that I have the right to withdraw either myself or any data provided by me without giving any reasons within up to three weeks after the data collection is completed.
- I understand that hard copy and digital data will be stored separately and securely for a period of six years and then destroyed.

- I understand that the data collected from the research will be used for the researcher's PhD thesis at the University of Auckland, and may be used for academic publications, and conference presentations, and if the information I provide is reported/published, confidentiality is assured and pseudonyms will be used to protect my identity.
- I know my confidentiality will be protected by the researcher and I will not be identified in any of the researcher's publications.
- I wish to receive a copy of the research findings by email \_\_\_\_\_ (If not, please leave this blank).

Name \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 DECEMBER 2017 FOR 3 YEARS, REFERENCE NUMBER 020446.