Libraries and Learning Services

University of Auckland Research Repository, ResearchSpace

Copyright Statement

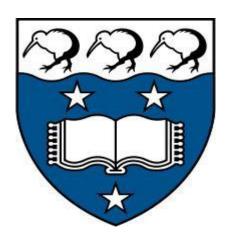The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.

Effects of Using Self-assessment on English-as-a-foreign-language (EFL) Students' Self-efficacy Beliefs and Writing Improvement

Xiaoyu Zhang

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (Ph.D.) in Education

(Applied Linguistics & TESOL)

The University of Auckland, Auckland

New Zealand

2021

ABSTRACT

Self-assessment and self-efficacy are two important strands in relation to instruction in EFL writing classrooms. Results from previous studies have generally suggested that self-assessment benefits writing performance, nurtures students' writing self-efficacy beliefs, and supports autonomous learning (Boud, 1995; Earl, 2013; Harris & Brown, 2018). Similarly, self-efficacy has been shown to be a strong predictor of academic achievement in the fields of education and educational psychology (Richardson, Abraham, & Bond, 2012). The close correlation between self-efficacy and self-assessment was also identified by many studies (e.g., Pajares & Valiante, 2006; Van Reybroeck, Penneman, Vidick, & Galand, 2017). However, the strong global endorsement of self-assessment and self-efficacy over the past two decades has not been realised in EFL writing classrooms, especially in the tertiary context of China. There is minimal research on how engagement in self-assessment affects Chinese undergraduate students' self-efficacy and writing performance.

The present study was designed to address the above-mentioned research gap by adopting quantitative and qualitative approaches, with the overarching purpose to implement a self-assessment-based intervention in Chinese tertiary EFL writing classes to foster confident, competent and autonomous EFL writers. Overall, 668 students and two lecturers from 15 medium to large scale universities participated in different parts of this study.

A total of 92 English major sophomore students from four intact classes and two English lecturers from a Chinese university participated in the main study for approximately four months. Two of the four classes formed the intervention group (51 students), and the

other two formed the comparison group (41 students). Students from the intervention group were provided with 16 weeks of self-assessment intervention, which was developed by the researcher, whereas those from the comparison group used peer-assessment in their usual English writing classes. Pre-and post-questionnaires, writing tasks, self-assessment of writing tasks, learning journals, semi-structured interviews, and class observation were utilised to collect data from student participants. Before and after the intervention, all students' self-efficacy for self-assessment of writing and writing performance were explored and the inter-relationships between these variables investigated. Students' perception of, and adaptation to, self-assessment of writing was also explored during the intervention to provide a better understanding of Chinese students' experiences in self-assessment of writing.

Data were analysed quantitatively and qualitatively to gain a clearer understanding of the implementation of self-assessment in EFL writing classrooms. The quantitative results show that Chinese EFL learners held a medium level of self-efficacy for self-assessment of writing, and their self-efficacy levels were enhanced after the self-assessment intervention. The results also indicate that, in the intervention group, there was a large increase in students' writing performance not only holistically but also in individual linguistic measures such as accuracy and fluency. In addition, both groups of students showed relatively lower rating accuracy in self-assessing individual writing dimensions compared with their overall writing performance. For the intervention group in the pre-test, although moderate to high degrees of correlation were found between students' self-efficacy for self-assessment of writing and dimensions of writing performance, around half of the correlations lost their statistical significance in the post-test. The qualitative findings reveal how students enhanced their self-efficacy for self-assessment of writing, writing performance, and rating accuracy throughout the

intervention. Additionally, the findings document students' perception of, and adaptation to, self-assessment of writing and the writing rubric. The factors that may influence students' self-assessment of writing practices, such as individual language proficiency, former teachers' feedback approaches, and cultural norms are also discussed.

This study contributes to research on self-assessment and self-efficacy in the EFL writing domain by yielding empirical evidence and pedagogical guidance for educators to embed self-assessment in their own contexts. In addition, by focusing on students' experience in self-assessment of writing, this study has generated insights into how students perceived self-assessment of writing and identified more effective approaches to conduct self-assessment. Theoretical and methodological contributions, practical implications, research limitations, as well as suggestions for future research are also discussed.

ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without many individuals' help and support in my Ph.D. journey. First and foremost, I am most obliged to my supervisors, Professor Lawrence Zhang, Professor Judy Parr, and Dr Christine Biebricher, for their guidance and support that inspired and sustained me over the years. They are the most wonderful supervisors anyone can ask for, and I know how privileged I am to have studied under their supervision.

I would like to express my deepest gratitude to my main supervisor, Professor Lawrence Zhang, for his continuous support and impeccable guidance throughout the passage of this doctorate journey. I have greatly benefited from his extensive knowledge on second language acquisition and statistical procedures. My sincere thanks also go to my co-supervisor, Professor Judy Parr, who always provided quick and incisive comments on my work at all stages although she had her own considerable workload. Her feedback on my drafts has helped me tremendously to grow as a researcher and writer. I would also like to thank Dr Christine Biebricher, whose advice and steadfast encouragement were invaluable to me to carry out this research.

I am much obliged to all the participants in my study without whom the study could not have been accomplished. I am also grateful to my best friend, Xin Zhang, who has provided me unwavering practical and emotional support throughout this journey. I am blessed to have her in my life.

Last of all, with heartfelt gratitude I dedicate this thesis to my family, for my parents' continuous love and unconditional support during this overseas venture in countless ways so that I could complete this thesis smoothly. I could not have asked for a better, more loving family and I will always be indebted to them for making me become the person I am today. A very special thank you goes to my husband, Weiyi, for his love, company and continuous efforts to make our lives better. I cannot thank him enough for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

There are three things extremely hard: Steel, a diamond, and to know one's self.

--Benjamin Franklin

1.1 Research Background

With the focus of higher education gradually shifting towards widening students' engagement in the learning process to enhance students' academic achievement, there is a growing recognition that the capacity to evaluate one's own work during studying or professional life is a vital contribution to individuals' success (Cassidy, 2007; Dochy, Segers, & Sluijsmans, 1999; Sambell, McDowell & Montgomery, 2013; Yan, 2020). Although challenging and time-consuming, self-assessment and peer-assessment, as more sustainable forms of assessment, are gradually gaining a more prominent role in higher education internationally (Boud, 1995; Boud & Soler, 2016; Earl & Katz, 2006; González-Betancor, Bolívar-Cruz, & Verano-Tacoronte, 2019).

The notion of self-assessment is not new as, in the 1930s, the benefits of self-assessment were already acknowledged by many authors in a range of disciplines and self-assessment is arguably a major strategy in developing students' assessment expertise and self-regulated learning skills (Panadero, Broadbent, Boud, & Lodge, 2019). Compared with external methods of assessment, self-assessment with sufficient training and assessment materials to develop students' self-assessment abilities, is believed to be a reliable approach to measure one's own language competence (Andrade & Brown, 2016; Bachman & Palmer, 1989; Birjandi & Siyyari, 2010; Oscarson, 2009; Ross, 1998;

Patri, 2002). Further, writing, as an important component of individuals' language competences, requires deliberate regulation of motivation and deployment of knowledge to be completed successfully (Hobson, 1996; Nielsen, 2012, 2019). Today, the heightened awareness of students' ownership of writing reinforces the necessity of practicing self-assessment regularly in every phase of writing (Boscolo & Hidi, 2007; Lee, 2017; Sadler, 1989).

Self-assessment is needed not only for students to become critical and efficient self-assessors of their writing but also to develop sophisticated levels of self-assessment to deal with unfamiliar professional situations successfully in the future (Bourke, 2014; Nielsen, 2019; Ross, Rolheiser, & Hogaboam-Gray, 1999; Sullivan & Hall, 1997). To realise these goals, in-time feedback on students' self-assessment efforts are needed to foster students' confidence in self-assessment and themselves (Harris & Brown, 2018; Xu, 2019).

1.2 Research Context

The present research was carried out in China, where English is taught as a foreign language from primary school (usually from Grade 3). The subsequent sections contextualise this research in English teaching and the classroom assessment practices in China.

1.2.1 Teaching English-as-a-foreign-language (EFL) writing in China

In an EFL context such as China, English writing abilities are essential for students to succeed in a variety of school entrance exams and professional qualifications, but university students are not intrinsically motivated to learn English writing (Ding & Zhao, 2019). For Chinese EFL learners, writing in English can be a challenging and daunting

task because of their insufficient knowledge of writing or writing strategies (Xiao & Chen, 2018). As a result, students may be demotivated and anxious in relation to writing (Yang & Gao, 2013).

Similarly, for Chinese English teachers, writing is a language area in which it is deemed difficult to develop students' competence (Gil, 2016; Sun & Wang, 2020; Zhang, 2013). There is little difference in writing instruction for English-major or non-English major students in Chinese universities; both focus on "micro-skills such as orthography and sentence level writing to macro skills such as paragraph and whole text writing" (Woodrow, 2011, p. 511).

Unlike non-English major students, English major students enroll in English courses specially designed to develop their English competence and knowledge in listening, reading, writing, and speaking as well as linguistics and literature (Zhang, 2016). According to the *National English Syllabus and Standards for English Majors* (Ministry of Education, 2018), English writing is a basic and compulsory course conducted in the second and third year study in universities for over four successive semesters. Although the duration of English writing courses varies in different universities, course content is consistent, covering various writing genres (e.g. exposition, summary, argumentation, and emails/letter writing), with an emphasis on models of grammar knowledge and text organisation (Li, Zhang, & Parr, 2020). Furthermore, the delivery of writing courses in EFL contexts is similar, meaning that those courses are usually teacher-centred and exam-oriented, without active engagement by students in the writing process.

The English-major students are expected to apply the writing skills they acquire in the classroom to succeed in the specially designed testing system, the Test for English Major (TEM) Bands 4/8, to measure their English language proficiency in writing, reading,

listening and translation. As a nationally acknowledged test, TEM Bands 4/8 is not only essential for evaluating the success of university English departments, but also critical for English-major undergraduates to succeed in future job hunting (Gil, 2016).

Chinese EFL writing teachers spend considerable time marking students' written work (Liu, 2002; Xiao & Chen, 2018), but, unfortunately, such marking and commenting leads to very few improvements in students' subsequent work. It appears that English-major students, even after studying linguistic constructions and different genres in the writing course, remain unclear about where they are and how well they are achieving relative to expectations (Hobson, 1996; Xu, 2019).

1.2.2 Classroom assessment practices in China

Regardless of the innovative practices in writing pedagogy in the past two decades, both teaching and learning of EFL writing in mainland China have long been exposed to an exam-driven, product-oriented, and teacher-centred learning environment (see Ng & Cheung, 2017 for a review). The forms of classroom assessment tend to be prescribed by authorities such as the faculty or administrative committees to meet the need of that environment (Chen & Zhang, 2019; Ding & Zhao, 2019; You, 2004). In such a situation, passing high-stakes examinations becomes the predominant motivation for Chinese writing teachers and students to perform English writing (Lee, 2014; Yu & Hu, 2017; Zhang, 2016). As there is little student involvement in the learning and assessment process (Carless, 2011; Xu, 2019), Chinese students tend to be demotivated because they are deprived of a sense of ownership and enjoyment in writing (Yao & Gao, 2013).

Unlike the increased number of studies in peer-assessment conducted in China (see Zhang & Cheng, 2020 for a review), self-assessment of writing seems particularly

challenging for Chinese learners. Self-assessment has not been widely used in China because teachers appear to have inadequate knowledge of self-assessment and they view it as an elusive and likely time-consuming construct (Lam, 2016, 2019). As Littlewood (1999) noted in his studies, East Asian students normally expect their teachers to be responsible for their learning assessment (Wang & Kim, 2017). Chinese students likewise tend to rely on their teachers and avoid direct engagement in the assessment process, because writing instruction focuses largely on test preparation at the cost of fostering learners' active role in assessing their learning. Students, therefore, are rarely informed about the process and the value of self-assessment or empowered to employ self-assessment with respect to their language knowledge and academic performance (Chen, 2008; Liu & Brantmeier, 2019; Panadero & Alonso-Tapiz, 2013).

However, as the ancient Chinese saying goes, teaching a person how to fish is better than giving the person a fish, in which 'giving a fish' could refer to teacher-directed classroom assessment practices that are fast but only effective in the short term. In contrast, 'teaching an individual how to fish' suggests equipping students with the skills and knowledge to conduct certain tasks (e.g., self-assessment) that could benefit students in the long term (Bai, Chao, & Wang, 2019; Yan, Brown, Lee & Qiu, 2019). With the Ministry of Education's (2018) call to position students at the heart of their learning, the potential benefit of self-assessment to respond to this call implies the need for current Chinese classroom assessment practices to transit from traditional assessment forms to student-centred assessment approaches.

1.3 Statement of Problems

This project in the field of self-assessment in EFL settings was driven by personal as well as academic research interests. On a personal level, learning more about myself as

an EFL learner, especially during writing, was pivotal; helping students to know themselves as language learners has also been central to my previous work as an English teacher. On a research level, however, the existing investigation into the knowledge and practices of self-assessment of writing is still preliminary and exploratory. It was surprising to find that studies concerning self-assessment are, arguably, still very limited. For instance, in western contexts, most research concerning student self-assessment centres on the elementary or the secondary level (e.g., Brown & Harris, 2013; Hawe & Parr, 2014; Ross et al., 1999), and there is scarce empirical research on self-assessment in tertiary writing classrooms.

Similarly, very few studies have been undertaken to date concerning self-assessment relevant to EFL teaching or practicing writing contexts in universities (see Butler & Lee, 2010; Liu & Xu, 2017; Matsuno, 2009; Wang, 2016), largely due to the difficulty for, and the reluctance of, both teachers and students to carry out systematic self-assessment (Carless, 2011; Earl, 2013; Earl & Katz, 2006; Panadero, Brown & Strijbos, 2016). For teachers, I believe this problem stems from an assumption that student self-assessment is challenging, impractical, or even not necessary (Oscarson, 1989; Xu, 2019) because teachers possess a limited understanding of self-assessment so that they are unable to guide students explicitly to attend to key aspects of self-assessment of writing (Earl & Katz, 2006; Lam, 2016). Teachers, furthermore, tend to have low expectations of students and therefore overlook students' potential to provide feedback on their own unless adequate support is given (Andrade & Boulay, 2003; González-Betancor, Bolívar-Cruz, & Verano-Tacoronte, 2019; Lee, 2011).

Students, it appears, cannot appropriately self-assess their writing because "they have not learned how to evaluate what they write; they have not internalized any consistent

set of criteria or standards to which they can hold themselves" (White, 1994, p. 10). Given the above situation, it is imperative for researchers and practitioners to investigate how to enhance EFL students' self-efficacy for self-assessment of writing and writing self-assessment competence.

These issues and the brief research background provided in section 1.1 highlight the following research areas that have received limited attention:

1) Investigation of the effects of using self-assessment in tertiary EFL writing settings on students' writing performance and linguistic development, especially exploring students' experiences in self-assessment of writing.

2) Assessment of learners' psychological variables, such as self-efficacy beliefs, in specific domains such as self-assessment of writing and exploring the interplay of self-assessment practices and learners' self-efficacy empirically.

3) Examination of the effect of using self-assessment on the development of students' rating accuracy from both quantitative and qualitative lenses.

Given these research gaps, I hypothesize that insights into self-assessment of writing practices in Chinese tertiary classrooms will advance an understanding of the role self-assessment plays in students' self-efficacy and writing performance.

1.4 Research Aims and Research Questions

Informed by constructivist, socio-cognitive, and formative assessment theories, this study investigated the effects of using self-assessment on EFL learners' self-efficacy for self-assessment of writing and writing performance. The integration of the three theoretical paradigms is expected to lead to a better understanding of students' self-assessment of writing in the EFL context.

7

To provide empirical evidence for the benefits of embedding self-assessment in the traditional tertiary level EFL writing classroom, this study had a threefold purpose. Firstly, the preparatory stage developed and validated, through factor analysis, a new instrument, *Self-efficacy for Self-assessment of EFL Writing* (SESAEFLW), to measure students' self-reported levels of self-efficacy regarding self-assessment of writing for use in the main part of this study. The SESAEFLW was constructed with multi-dimensions of self-efficacy for self-assessment pertinent to writing skills and writing tasks. Secondly, Part One of the main study examined whether the use of self-assessment affected students' self-efficacy beliefs and writing performance, as measured by indexes such as complexity, fluency, and accuracy; data were collected from self-reported questionnaires, two writing tests, two self-assessment of writing tasks, and classroom observations. Thirdly, Part Two of the study explored how students, with different English proficiency levels, perceived and reacted to self-assessment of writing, as well as to the writing rubric used in the self-assessment. Potential factors that could affect students' self-assessment practices, such as English proficiency and the feedback approach in previous education, were also considered in Part Two. Overall, the aim of this study was to document the findings as a basis for further deliberation on self-assessment in higher education and to contribute to the existing literature on self-assessment of writing practices and pedagogical innovation in the EFL writing context.

The following research questions directed this study ('students' refer to Chinese undergraduate English-major students).

1. What effects does the use of self-assessment in the EFL writing class have on:

    a) Students' self-efficacy for self-assessment of EFL writing?

b) Students' writing performance in different dimensions as well as their linguistic improvements in terms of complexity, accuracy, and fluency?

c) Students' rating accuracy in self-assessment of writing as related to the raters' assessment?

2. How does students' self-reported efficacy for self-assessment of writing correlate with their writing performance?

3. What are students' (e.g., basic/proficient/advanced English proficiency) perceptions of, and attitudes to, self-assessment of writing, and the rubric used for self-assessment of writing?

4. What factors influence EFL students' self-assessment of writing practices?

1.5 Significance of the Research

This research has the potential to contribute to the following aspects of the field. Theoretically, this study is expected to extend understandings of an aspect of foreign language acquisition, the use of self-assessment in the development of EFL writing, through the lenses of constructivism, self-regulated learning theory, and formative assessment theory (Harris & Brown, 2018). It is also postulated that, drawing on those theories, a self-assessment-based intervention may enhance current classroom assessment practices in China through student-centred assessment approaches. This study is also expected to contribute to understanding how self-assessment of writing practices could influence learners' psychological factors (such as self-efficacy) and writing performance in an EFL context.

The results of this study, using a quasi-experimental approach will provide much needed empirical quantitative evidence of the potential value of student-centred formative

assessment, especially self-assessment in the EFL writing context (Carless, 2011; Lam, 2019; Liu & Xu, 2017). Furthermore, the current study contributes to research on self-assessment in the EFL contexts by identifying the effects of self-assessment on the linguistic features of writing quality: complexity, accuracy, and fluency to complement regular measurements in the writing rubric (Lu, 2017). This study also reveals the complex nature of Chinese students' experiences in self-assessment, which are influenced by individual and cultural factors.

The current study contributes to the field, methodologically, through the development and validation of the *Self-efficacy for Self-assessment of EFL Writing* questionnaire. This newly developed research instrument has the potential to be employed in other contexts. Methodologically, this study also contributes through the research design, which has the potential to inform, and to be used for, valid and in-depth investigations of self-assessment in similar EFL writing settings. The main study, which consisted of two complementary parts, used multiple research methods (quantitative and qualitative) for data triangulation.

Pedagogically, the present study provides empirical evidence for EFL writing teachers' classroom instruction, in terms of developing active and capable self-assessors for effective learning. The inclusion of self-efficacy in the current study has the potential to enhance the feasibility and effectiveness of using self-assessment-based intervention for Chinese EFL learners' writing development. The pedagogic innovation in using a student-centred assessment method, self-assessment, in a regular writing curriculum, may alleviate EFL writing teachers' assessment tasks so that more energy is available for other effective classroom practices (Lam, 2019; Liu, 2002; Liu & Xu, 2017; Oscarson, 1989). Empirical evidence from the self-assessment of writing practices is

expected not only to increase teachers' willingness to implement self-assessment but also to raise students' awareness of their writing progress and quality. Furthermore, teachers and students are expected to establish a stronger collaborative relationship through the self-assessment process (Shepard, 2000).

Practically, EFL students will have a deeper understanding of their academic selves, their writing efficacy, and the criteria for quality writing through the constant engagement in self-assessing their writing products (Bourke, 2014; Hobson, 1996; Usher & Pajares, 2008). For instance, the questionnaire designed for data collection and analysis in this study can be used to raise students' consciousness in self-assessment of writing. However, the most significant point of embedding self-assessment in the EFL writing classroom is its alignment with the practical expectations of the globalised world, and the changing goals of higher education. As Boud (1995) argued persuasively, "the ability to self-assess is a key foundation to a career as a lifelong learner who can continue their education after formal education has ended" (1995, p. 14). To be prepared for life beyond the University, students are expected to: Accept responsibility for their own learning; exhibit transferable skills such as independence, self-management, and critical-reflection, and to carry on their further education or professional development without external assistance to increase their employability (Dochy et al., 1999; Gil, 2016; Sambell et al., 2013). This study, therefore, may be a useful contribution through encouraging students to become independent learners, using self-assessment of writing as a starting point.

1.6 Organisation of the Thesis

This thesis comprises eight chapters. Chapter One describes the research concerning its background, context, existing problems, research purposes and questions, as well as the

significance of the research holistically. Following this introductory chapter, the second chapter presents a discussion of prior relevant studies that have informed this research focusing on self-assessment and self-efficacy with respect to EFL writing, particularly in the context of China. Chapter Three establishes the theoretical foundation of this study as constructivism, self-regulation, and formative assessment theories to show that self-assessment can be understood via multidimensional lenses. As well as an explicit description of the instrument development and validation process, Chapter Four describes the research methodology adopted.

Chapters Five and Six present the findings, reporting the effects of the self-assessment-based intervention on students' self-efficacy and writing performance. Specifically, Chapter Five presents the quantitative results of students' perceived self-efficacy for self-assessment of writing, students' writing performance and linguistic development, and the relationships of those variables. Students' rating accuracy, compared with the raters' rating is also presented in this chapter. Chapter Six provides the qualitative data documenting how six students, with varied English proficiency levels, navigated their way through self-assessment of writing. Chapter Seven presents a discussion integrating the research findings and viewing them in relation to existing theory and literature. The concluding chapter, Chapter Eight, first provides a brief review of the current research findings, and then presents the theoretical and methodological contributions, as well as the pedagogical implications. Chapter Eight also concludes the thesis identifying the limitations of this research and suggestions for future research.

CHAPTER TWO

LITERATURE REVIEW

This chapter reviews the previous relevant literature that has motivated and generated the research questions for this study. The literature review starts with the recent developments in assessment and learning approaches, together with an introduction of the key concepts in this study. Two areas of research concerning self-assessment are then presented: formative assessment and self-regulated learning [SRL] (Andrade & Brookhart, 2016). Within theories of formative assessment, self-assessment is used as a pedagogical tool in teachers' instructional approach. Teachers' role in student self-assessment and issues and challenges faced by self-assessment implementation, therefore, are considered in this chapter. When situating self-assessment within SRL, self-assessment focuses on learners' cognitive level and it is understood as "a process that pupils carry out to self-regulate their learning" (Panadero & Alonso-Tapia, 2013, p. 554). The influence of using self-assessment on students' learning achievement and the application of self-assessment in the writing context, in the context of SRL, are presented in detail. The ensuing sections evaluate empirical studies assessing writing self-efficacy and the interplay between self-efficacy and self-assessment. Finally, a summary of the research gaps identified in this review of the literature, that the present study aims to address, is presented.

2.1 Self-assessment and Social Cognitive Theory

The ever-changing interplay and alignment of assessment and learning have attracted interest amongst researchers and educationalists alike in the past two decades to develop a better understanding of how assessment can contribute to students' learning (Bouziane

& Zyad, 2018). Various terms have been employed over the years to indicate different assessment and learning approaches, for example, formative assessment, summative assessment, assessment for learning (AfL), assessment of learning (AoL), assessment to learn (AtoL), and assessment as learning (AaL).

Based on the functions of classroom assessment, formative assessment and summative assessment were proposed (Bloom, Hasting & Madaus, 1971). Black and Wiliam (1998) defined formative assessment as "encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (p. 2). In contrast to the ongoing nature of formative assessment, summative assessment tends to be conducted at the end of a given period through timed examinations providing judgement on students' learning accomplishment by marks or grades (Lam, 2016).

Later, the concept of AfL was introduced by the UK Assessment Reform Group (ARG) in 1999, and has attracted scholars' interest globally. Broadfoot et al. (1999) proposed to use the term AfL instead of formative assessment as the term 'formative' was not sufficiently specific, giving rise to a wide range of interpretations. The definition of AfL as "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (ARG, 2002, p. 2) is widely accepted. This definition was later updated stressing the dynamic nature of AfL, indicating that it "is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance on-going learning" (Klenowski, 2009, p. 264).

The terms formative assessment and AfL are not synonymous. According to Black, Harrison, Lee, Marshall and Wiliam (2004) who have differentiated them, the first priority of AfL is to promote students' learning, whereas the purpose of formative assessment is to adapt the teaching to fulfil students' learning needs. Many authors still tend to use those terms interchangeably in their work. For example, it is commonly believed that AfL pertains to formative assessment as the central purpose of formative assessment maintaining a focus on students' engagement in assessment, while AoL is equated with summative assessment because it emphasises more on students' grades and outcomes (Laveault & Allal, 2016).

A growing need for students to participate actively in assessment practice in the school context has blurred the boundaries between assessment and learning. The concept of Assessment as Learning (AaL) has emerged (Dann, 2002, 2014; Earl, 2013) and has then been addressed in a number of publications from different perspectives (e.g., Lam, 2016; Torrance, 2007, 2012). Based on AfL, the origin of the concept AaL was promoted in Dann's (2002) study, in which AaL is arguably, "not merely an adjunct to teaching and learning but a process through which pupil involvement in assessment can feature as part of learning" (p. 153). In other words, AaL assigns the role of the assessor to students themselves in the assessment processes and expects them to be capable of supporting and advancing their own learning (Earl, 2013). Thus, the concept of AaL encompasses the complex interplay of assessment, teaching, and learning (Dann, 2014).

Torrance (2007) posited that assessment has moved from AoL, through to the widely adopted notion of AfL, to AaL. Such a trend means that students' experiences and involvement in assessment should be optimised and regarded as indispensable to advance their learning. Because of the inadequacies of traditional teacher assessment,

AaL further directs scholars' attention to alternative or diversified forms of assessment, such as self-assessment, which is at the heart of AaL. AaL is of great value in giving the ownership of learning back to students (Andrade, 2019; Ross, Rolheiser, & Hogaboam-Grey, 1999), especially in the teacher-driven, English as a foreign language (EFL) writing contexts (Lee, 2011, 2017). It would seem that all the efforts learners put into other assessment practices, ultimately, contribute to their capability in self-assessing (Earl & Katz, 2006; Lam, 2016).

Although the field has developed, as discussed, there are issues that still need to be addressed. For instance, scholars' endorsement of self-assessment has not made such practices commonplace in educational settings, and there is a lack of a concrete definition of self-assessment in either assessment or learning contexts. The following section examines and discusses definitions of self-assessment in previous studies.

2.1.1 Definitions of self-assessment

In the recent review (Panadero et al., 2016), self-assessment appears a rather nebulous term as more than 20 different categories of its implementations were found. Unsurprisingly, there is no uniform definition of standard self-assessment due to its complexity and multifaceted nature, and the range of activities it embraces (Andrade, 2019). A widely accepted definition of self-assessment indicates that it applies whenever learners are involved in making "judgements about their own learning, particularly about their achievements and the outcomes of their learning" (Boud & Falchikov, 1989, p. 529). Similarly, self-assessment is also referred to as a "wide variety of mechanisms and techniques through which students describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products" (Panadero, Brown, Strijbos, 2016, p. 804). In some definitions, a learning-oriented

purpose is added to self-assessment (Andrade, 2010; 2019; Yan, 2018), which suggests that self-feedback generation and performance improvements should also be included when defining self-assessment (Huang, 2016).

Many scholars have argued that self-assessment should not be used when scoring or grading one's own work is involved; a distinction has often been made between self-assessment and self-evaluation (Andrade, 2010; Andrade & Du, 2007; Andrade & Valtcheva, 2009; Panadero et al., 2016). For example, self-assessment indicates a formative process in which students deeply reflect upon their learning and work quality while self-evaluation implies students making summative judgements about their work that result in a final grade or mark (Andrade & Du, 2007). Some researchers have argued, however, that before making more realistic and sophisticated self-judgements, students need to develop their self-assessment competence progressively, and that simple self-scoring or grading practices (summative judgements) could be a useful starting point (Brown & Harris, 2014; Harris & Brown, 2018).

Echoing Oscarson's (1989) proposal of two types of self-assessment (performance-oriented and development-oriented), Tan (2007) further considered self-assessment from three different perspectives according to the distribution of power between teachers and students as well as the degree of students' contribution, during the self-assessment process. Although all three aspects, teacher-driven, programme-driven, and future-driven, could contribute to the growth of student self-assessment skills, more importance is attached to the future-driven student self-assessment as it emphasises students' motivational development, and sustainability of self-assessment capacity beyond the limits of course content and requirements (Bourke, 2014).

To avoid terminological confusion, throughout the present study, I use the concept self-assessment as an umbrella term, under which terms such as self-marking, self-revision, and self-evaluation are included (Sadek, 2018). In this sense, self-assessment focuses on both learning processes and learning results in two forms, namely, formative and summative (Panadero, Brown, & Strijbos, 2016; Van Reybroeck, Penneman, Vidick, & Galand, 2017). In the formative sense, self-assessment is "feedback for oneself from oneself" (Andrade & Du, 2007, p. 160), and refers to a self-regulatory process that learners create and experience evaluation concurrently (Henner-Stanchina & Holec, 1985; Panadero & Alonso-Tapia, 2013; Paris & Paris, 2001). In that sense, self-assessment can be conceptualized as a cyclical process, "during which students collect information about their own performance, evaluate and reflect on the quality of their learning process model and outcomes according to selected criteria, to identify their own strengths and weaknesses." (Yan & Brown, 2017, p. 2). Students, first, gather, monitor and reflect on information about their knowledge, performance, and attainment while learning; they then identify and evaluate the potential strategies to improve aspects of their learning against personal, or pre-established, criteria to revise their work accordingly (Andrade & Boulay, 2003; Andrade, Du, & Mycek, 2010; Black & Wiliam, 1998; Boud, 1995; Klenowski, 1995; Longhurst & Norton, 1997; McMillan & Hearn, 2008; Yan, 2018).

The summative assessment indicates that language teachers apply self-assessment as a measurement tool to empower and encourage students to reflect on and improve their performance (Boud, 1995; Butler & Lee, 2010; Harris & Brown, 2018). In this sense, probably in the majority of the current self-assessment practices, learners simply rate, or grade, their final work to understand their proficiency in a certain task (Boud, 1995; Pinner, 2016). While summative self-assessment practices pressure students to make a

judgement (Boud, 1995), they can be used also to supplement formative self-assessment, which includes a wide range of exercises such as self-revising/reflection, performance estimation, criteria-or rubric-based assessments (Andrade & Valtcheva, 2009; Nielsen, 2019).

2.2 Positioning Self-assessment as a Pedagogical Tool within Formative Assessment Perspectives

Formative assessment has been actively promoted by the United Kingdom's Assessment Reform Group (ARG, 2002; Black, Harrison, Lee, Marshall, & Wiliam, 2003; Black & Wiliam, 1998) because of its potential in enhancing students' learning within the global assessment for learning movement. Perhaps because formative assessment endorses a plethora of practices, a variety of opinions still exist in defining the concept of formative assessment. There appears, however, to be agreement that the purpose of the assessment approaches and effects makes the assessment formative or not (Sadler, 1989). Since students' active participation in, and use of, assessment is a salient feature of formative practices, formative uses of self-assessment, for example, students use self-assessment to generate feedback for themselves before formal grading, are promoted as a valuable form of sustainable assessment (Andrade, 2010; Black et al., 2003; Boud & Soler, 2016; Harris & Brown, 2018).

From a pedagogical perspective, particularly at the tertiary level, self-assessment has been promoted as a valuable learning strategy that teachers could draw on, generating formative feedback, to improve students' learning outcomes (Bourke, 2014; Xu, 2019). Similarly, when teachers encourage students to reflect on their work, they not only have the opportunity to utilise feedback-to-self- to revise their work prior to formal evaluation (Panadero & Alonso-Tapia, 2013), but also develop their evaluative capability (Tai,

Ajjawi, Boud, Dawson, Panadero, 2018). The following sections describe the role of teachers in student self-assessment and the issues for its effective implementation.

2.2.1 Teachers' role in student self-assessment

Apart from emphasising students' agency in the process of self-assessment, the changed role of teachers in students' self-assessment also needs to be considered (Gardner & Miller, 1999; Lee, 2011). Firstly, teachers, rather than dominating the assessment process, need to relinquish control over students and understand that sharing and distributing their power to students during self-assessment is beneficial (Tan, 2009), and that their mediating and modelling role is important in engaging students in the self-assessment process (Dann, 2014; Panadero et al., 2016). Secondly, as well as conveying relevant knowledge to students (Sadler, 1989), teachers need to adopt multiple roles such as the counsellor, materials developer, administrator and organiser in the classroom (Lee, 2011, 2017). Their roles, however, are not limited to these; for instance, teachers should uphold students' dignity when they turn to teachers for feedback concerning their self-assessment. Students need to know that teachers are confident in their ability to assess their own writing, so that their self-efficacy in relevant tasks is enhanced (Schunk & Usher, 2011).

The following section reviews how teachers, given the complexity of their roles in self-assessment and the skills required, can integrate self-assessment in the classroom, scaffold students in self-assessment procedures, and strengthen the positive effects of self-assessment.

2.2.1.1 Scaffolding and implementation of self-assessment

Scaffolding students' self-assessment and teaching them the regular content and skills at the same time is challenging for teachers, but these are the requirements for students to generate accurate self-assessment (Andrade & Brown, 2016). It is argued that teachers' scaffolding is like explaining to students the learning criteria "from latent to manifest and back to latent again" (Sadler, 1989, p. 134), which implies "guiding learners towards greater autonomy by gradually removing teacher support" (Woodrow, 2011, p. 520). "A fine-tuned self-assessment ability does not come automatically to all students" (LeBlanc & Painchaud, 1985, p. 675), which implies the need for formal and systematic practices for students to internalize self-assessment, with scaffolding by teachers. Teachers' scaffolding is essential for students' self-assessment development (Earl & Katz, 2006), especially when "there is clear evidence shown that the learner is unable to progress without assistance" (Clark, 2012, p. 230) or students' have inadequate domain-specific knowledge and skills (Panadero, Jonsson, et al., 2016). In the classroom context, scaffolding for self-assessment occurs in many ways, for instance, awareness-raising, provision of guidance and learning materials, and interpretation of results (Gardner, 2000; Ross, 2006).

Other researchers also identified similar means to scaffold self-assessment, through preparing students' for self-assessment, stimulating students' active and authentic involvement, and providing practice with feedback in a trustworthy environment (Andrade & Brown, 2016; Harris & Brown, 2018; Rolheiser, 1998; Sambell et al., 2013). Successful scaffolding in a writing context can involve discussion, negotiation or co-construction of self-assessment criteria with students prior to, during, or after writing so that students are familiar with the specific domain (Lee, 2017). The specific approaches

that teachers can employ to scaffold students' practicing of self-assessment and the time that teachers need to remove such scaffolding, however, are still not clear (Zhang & Zhang, 2018).

Rolheiser (1996) proposed a three-level model in four stages (as shown in Table 2.1) to implement self-assessment, in which the teacher acts as a coach and a model of self-assessment practices (Earl & Katz, 2006; Earl, 2013). Students are engaged in different stages of self-assessment with the teacher transferring, gradually, more responsibility and freedom to students, to enable students to focus on the quality of their work rather than a grade (see Table 2.1). Consistent with Rolheiser's (1996) self-assessment implementation model, Andrade and Du (2007) framed the self-assessment process as a three-step pedagogical procedure with increasingly less teacher involvement, in which, initially, expectations concerning learning in the particular area are shared between teachers and students. Secondly, students embark on their assignments and check them against the teacher-provided rubric, following which, students revise and improve their work based on the self-generated feedback.

As well as the crucial role of students' internal feedback in self-assessment, it can be seen from Table 2.1 that teachers' feedback is equally important across all levels. For example, at the beginning level, with detailed feedback from the teacher, students could progress their learning. Students should then be taught how to interpret feedback, make connections between the feedback and the work they produce, and apply feedback to improve future work (Andrade & Valtcheve, 2009; Sadler, 1989). To promote willing and honest self-assessment practices at the initial stage, it is suggested that teachers allow students to keep their self-assessment partially private (Andrade & Brown, 2016; Harris & Brown, 2018).

Table 2. 1 Teacher's Role in Implementing Self-assessment in Four Stages.

	Stage 1	Stage 2	Stage 3	Stage 4
Levels of Imple-mentation	*Raising Awareness and Establishing Criteria*	*Teaching Students How to Apply Criteria*	*Providing Feedback to Students on Application of Criteria*	*Setting Learning Goals and Strategies*
Beginning	Criteria are given to students for their reaction and discussion	Examples of applying criteria given to students	Teacher provides feedback	Goals and strategies determined by the teacher
Inter-mediate	Students select criteria from a menu of possibilities	The teacher describes and models how to apply criteria	Feedback provided by both teachers and students	A menu of goals and strategies is provided by the teacher
Advanced	Students generate criteria on their own or with the teacher	Students' apply criteria to their own work	Students initiate and justify their' own feedback with the teacher's help	The student constructs goals and strategies to achieve the goals with the teacher's guidance

Note. Adapted from Rolheiser (1996)

At the advanced level of implementation, the feedback, which could focus on either students' work or students' self-assessment (Harris & Brown, 2018), has to be so "transparent that students can evaluate their own work in the same way that their teacher would" (Shepard, 2000, p. 11). The timing and the involvement of encouragement in teachers' feedback are pivotal in all levels of self-assessment, as students need sufficient time to digest teachers' feedback for further actions and students need external assurance to enhance their confidence in practicing self-assessment (Chen, 2008; Harris & Brown, 2018; Panadero et al., 2016). Lee (2017) further indicated three useful principles for teachers to inform their feedback; they can be summarised as a selective and personalised response, of a diagnostic and concrete nature, that is encouraging and

integrates with learning. Such feedback could support students in assessing their own knowledge and performance and in using self-feedback to overcome the instincts of ego protection and avoiding possible ego-inflation resulting in students' overestimation (Earl & Katz, 2006; Pintrich, 2002).

Pivotal to the successful implementation of scaffolding is for teachers to provide students with an explicit explanation of why, how, and when to self-assess, followed by relevant activities and well-timed feedback to support students' realistic self-assessment (Harris & Brown, 2018).

2.2.2 Issues for effective implementation of self-assessment

The following sections elaborate and discuss the issues and potential challenges of self-assessment implementation reported in previous studies.

2.2.2.1 Fidelity of student self-assessment

Previous studies have used a range of terms, such as accuracy, consistency, reliability, and validity, when discussing the alignment of students' and teachers' judgements in the complex field of self-assessment. As false self-assessment could possibly lead to students' inaccurate understandings of their abilities (Harris & Brown, 2018), alignment, and curricular expectations in the assessment process, in the writing domain, have been a focus of research in EFL writing in the past two decades (Sadek, 2018; Zheng & Yu, 2019).

There is consensus that accuracy equates with validity, and consistency with reliability (e.g., Brown & Harris, 2013; Earl & Katz, 2006; Harris & Brown, 2018; Ross, 2006; Weigle, 2002), however, there is confusion with terminology when accuracy and

consistency are sometimes used interchangeably (Brown, Andrade, & Chen, 2015). In a recent study for example, when discussing the alignment between students' and external raters' evaluations, the "accuracy of self-assessment" was considered a less accurate term compared with the "consistency of self-assessment" (Andrade, 2019, p. 5).

Underpinning the definitions of these terms is the need to indicate the truthfulness, or the reality, of students' self-assessment in comparison with the external assessment. The present study, therefore, proposes the fidelity of self-assessment as an umbrella term to represent the general quality of student self-assessment, so that educators and learners may find it easier when applying self-assessment in their own contexts. To avoid confusion, however, I have maintained the specific term other scholars have used in their work when reviewing the studies that follow.

Although self-assessment could be a useful pedagogical practice in most educational contexts, when measuring students' competence, due to its subjective nature as a measurement tool, it has been often questioned (Butler & Lee, 2010; Harris & Brown, 2018), and contradictory findings reported in the literature concerning the validity of students' self-assessment (Bachman & Palmer, 1989; Blue, 1988; Boud & Falchikov, 1989; Brown et al., 2015; Brown & Harris, 2013; Leach, 2012; Oscarson, 1989; Patri, 2002). For instance, a lack of congruence has been reported between students' self-assessment scores and those of the teachers', and students' self-assessment marks were not found to be comparable to their actual academic abilities (Boud, 1989; González-Betancor et al., 2019; Kirby & Downs, 2007; Nawas, 2020; Orsmond, Merry, & Reiling, 1997). Other studies also found that there is a tendency for high achievers to under rate and low achievers to over rate themselves (Boud & Falchikov, 1989; Kun, 2016; Leach, 2012; Stefani, 1994; Sullivan & Hall, 1997). The fact that students are inclined to

overestimate or underestimate their work makes the validity of self-assessment open to doubt (Harris & Brown, 2018; Matsuno, 2009; Topping, 2003). The problem, however, might be in the implementation of the self-assessment, and applying the self-assessment outcomes in a summative manner (Brown & Harris, 2014; Leach, 2012), such as when using students' self-marking mainly to substitute for teacher-marking (Tai et al., 2018).

Social and psychological factors may also affect students' self-assessment validity; for instance, Blanche and Merino (1989) in their comprehensive review, indicated that students' self-assessment accuracy could be affected by a series of factors, and largely depends on students' linguistic proficiency and the resources they use during the assessment (Kun, 2016; Topping, 2003). Other researchers also pointed out that over optimism about one's academic ability or effort, lack of vital contextual information, and classroom pressure to promote self-value (Dunning, Heath, & Suls, 2004) could lead to inaccurate outcomes of students' self-assessment. Further, students' age, cultural background, experience with the rating rubric, and the complexity of the targeted task should be considered when evaluating the validity of students' self-assessment (Brown & Harris, 2013; J. Zhang, 2016).

Among the above possible factors, students' cultural backgrounds have been shown to be a significant negative impact on their self-assessment validity. It is argued that some nationalities tend to overrate their language proficiency, whereas others are inclined to undervalue it (Blue, 1988). For example, in Asian countries, being modest about one's ability is highly valued (Chen, 2008; Matsuno, 2009), and so giving oneself a good assessment may be interpreted as boasting (Brown & Harris, 2013). Whether self-assessment is performed privately or publicly may also impact the validity of students' self-assessment (Harris & Brown, 2018).

Contrary to literature that suggests a lack of alignment between student self-assessment and teacher assessment, other studies have reported relatively good levels of agreement between students' and teachers' grading (Boud & Falchikov, 1989; Leach, 2012; Mazloomi & Khabiri, 2018; Stefani, 1994; Sullivan & Hall, 1997), or at least weakly positive correlations between students' and teachers' judgements (Brown & Harris, 2013). For example, in Leach's (2012) study, when university students ($n = 472$) were provided the opportunity to self-assess their work, the results showed highly significant statistical correlations between students' and the teacher's rating. Nevertheless, it is argued that using teacher judgement as the standard to evaluate students' self-assessment validity is problematic, in part because teachers and students may have different sources of information and knowledge bases for assessment (Andrade & Brown, 2016; J. Zhang, 2016). Therefore, the above factors and results should only serve as aspects to consider in students' self-assessment practices, rather than the basis on which to reject self-assessment (Gardner, 2000) because the positive role self-assessment can play in learning outweighs this possible lack of validity.

It seems the reliability (consistency) of student self-assessment has not received attention equal to that given to its validity. The reliability of students' self-assessment, however, is similarly problematic as reflected in the inconsistent findings reported in the literature (Earl & Katz, 2006; Meihami & Razmjoo, 2016; Ross, 2006). Prior studies have noted that students, especially mature students, can make judgements consistent with their previous self-assessments (Ross, 2006; Xu, 2019); but such consistency across tasks, items and subjects is apparent only when self-assessment is performed over short time periods, not over a longer time frame (Meihami & Razmjoo, 2016). Further, even if systematic self-assessment training is provided, the reliability of self-assessment remains affected by a range of factors, such as the amount of scaffolding, practice and

feedback that teachers provide (Panadero, Brown, et al., 2016; Topping, 2003), and the confidence students have in a series of metacognitive process such as self-reflection and self-modification (Earl & Katz, 2006).

Given that the validity and reliability of student self-assessment could be influenced by various factors, previous studies have discussed effective approaches of classroom self-assessment application. For instance, it is argued that the effectiveness of self-assessment should not be judged by the accuracy of students' self-assessment or similarity of students' and teachers' marking; it should be judged based on students' development during the assessment process (Chen, 2008; Orsmond et al., 1997). Moreover, academics may need to place more emphasis on students' self-assessment reflexivity and sustainability rather than being concerned about its reliability (Tan, 2007). Therefore, to maximise the fidelity of self-assessment, greater attention should be paid to how students' learn and how teachers support students' self-assessment rather than comparing students' self-generated grade with the external assessment (Brown et al., 2015; Butler, 2018; Yan & Brown, 2017).

2.2.2.2 Challenges of self-assessment implementation

While self-assessment may be easier for students since it does not confront them with the social difficulties that they may have when assessing their peers' work, a number of research studies have described a range of challenges with self-assessment for both educators and students (Tai et al., 2018). For educators, scaffolding students' self-assessment and teaching them the regular content and skills at the same time is a vital challenge (Andrade, 2010). From the EFL students' point of view, a major challenge of self-assessment is that English is a foreign language for them, in which the script, syllabus and grammar could be different and difficult for them (Burner, 2016).

Previous studies have identified other challenges for self-assessment, for example, Meihami and Razmjoo (2016) discussed several challenges that were perceived by both teachers and students when employing self-assessment in tertiary EFL writing courses. The first challenge is that the impact of cultural values, for instance, modesty and self-confidence, may impede self-assessment in certain contexts and constrain students being honest in self-assessment. Some cultures place more value on the summative purposes of self-assessment, which may prevent self-assessment from impacting on students' learning progress if it is the only form of self-assessment implemented (Harris & Brown, 2018; Lee, 2011b).

The second challenge concerns teachers' reluctance to promote self-assessment, even when they are offered the opportunity (Hawe & Parr, 2014; Lee, 2017), because of their lack of assessment literacy, including a limited understanding of assessment purposes and value, and limited experience in using effective student-centred practices. If teachers implement self-assessment without sufficient assessment knowledge, it might be superficial and not lead to students' improvement (Panadero et al, 2016). Furthermore, teachers' lack of assessment literacy is associated with whether they believe students could conduct self-assessment effectively, and whether teachers are able to work collaboratively with their colleagues to design context-specific self-assessment practices (Andrade & Brookhart, 2016; Burner, 2016).

The next challenge is the time-consuming nature of self-assessment. Students need considerable time to practice and to become accustomed to self-assessment. Likewise, teachers need time to provide personal feedback during self-assessment, considering students' sense of vulnerability, particularly with a large number of learners in L2 classrooms (Lee, 2011b; Liu & Xu, 2017). For EFL practitioners, there is a challenge in

achieving a balance between affording students sufficient independence to experience a sense of control, and offering them sufficient guidance and facilitation in self-assessment writing procedures to achieve their goals (Zhang, 2016). Integrating self-assessment into teaching practice, therefore, may increase teachers' workload and working time substantially, especially in the early stages of implementation (Hanrahan & Isaacs, 2001), yet allowing time may be essential to practice self-assessment of writing effectively. It is not clear how teachers can engage in dialogue about self-assessment of writing, and include students, within the limited time. Given the challenges, self-assessment activities may be a waste of classroom time, without advancing students' learning, if they are not well designed (Panadero et al., 2016).

As students' lack of capability and their resistance to self-assessment practices are common, another challenge is developing a psychologically congenial and supportive learning atmosphere that frees students from their long-standing disempowerment in the traditional assessment activities (Liu & Xu, 2017; Tan, 2004, 2009). In such an environment, students are not afraid of making mistakes, instead, students regard their mistakes as a normal part of learning (Earl & Katz, 2006), and can generate realistic self-assessment, which separates judgement of a task from self-worth, and share truthful feelings and feedback with their teachers (Harris & Brown, 2018; Sambell et al., 2013). Challenges when implementing self-assessment are multiple, and approaches to meet those challenges are equally diverse. The goal, however, is to ensure that teachers are well supported to help students develop the knowledge and capabilities needed for effective engagement in the self-assessment process.

2.3 Locating Self-assessment within Self-regulated Learning (SRL) Theories

In educational research, self-assessment and SRL are two salient tools to empower learners to organise and to monitor their learning (Andrade, 2019; Nicol & McFarlane-Dick, 2006; Sambell et al., 2013; Xu, 2019). Conceptual and practical overlaps between self-assessment and SRL have been acknowledged widely (Andrade & Brookhart, 2016; Andrade & Valtcheva, 2009; Brown & Harris, 2013; Panadero, Jonsson, & Strijbos, 2016; Paris & Paris, 2001) and have spawned extensive empirical research (see review by Panadero, Andrade, & Brookhart, 2018).

Self-assessment and self-regulated learning, as the word 'self' implies, both emphasise the proactive role of learners in setting goals, deploying and regulating actions and providing feedback to self to improve their academic performance (Andrade, 2010; Zimmerman, 2008). According to Pintrich (2002), SRL is "an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment" (p. 453). SRL, thus, can be seen as the formative use of self-assessment (Andrade, 2010, Liu & Xu, 2017).

Although self-assessment and SRL are inextricably linked, when self-assessment is interpreted differently, for example, as a learning strategy or a process, the relationship may vary. For example, if self-assessment is considered as a learning strategy for students' to judge their learning product, then self-assessment is an essential part of SRL skills as research on self-regulated learning has tended to encompass both processes and product (Brown & Harris, 2014; Panadero et al., 2019). Self-assessment, viewed as a process, is similar to SRL, because it not only comprises "all three domains of self-regulated learning: cognitive, motivational and affective" (Paris & Paris, 2001, p. 95),

but also supports and influences all stages of the self-regulatory cycle. Self-assessment is a process, therefore, to assist students in learning to self-regulate their thoughts, feelings and behaviours with valuable self-feedback (Harris & Brown, 2018; Panadero & Alonso-Tapia, 2013; Panadero et al., 2016; Yan, 2020).

The model in Figure 2.2 (Harris & Brown, 2018, p. 21) provides details on how self-assessment functions throughout SRL processes. The dashed lines indicated students' self-assessing process, in which students' internal competence, thoughts and beliefs interact with the comments they achieve from parents, teachers and/or peers. Through retrospective self-assessment of learning products and processes, it is anticipated that students will improve their learning outcomes and capabilities.
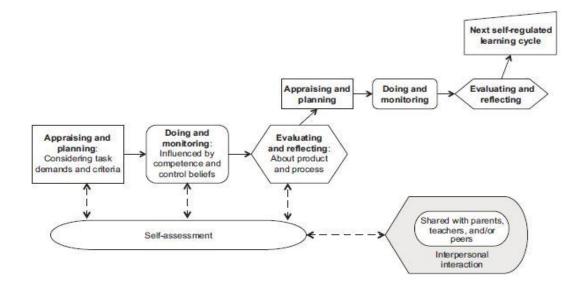
Figure 2. 1 Self-assessment in the SRL processes

Note. From Harris and Brown (2018, p. 21)

Self-assessment seen as a process from a phase-based point of view, with its aims and phases largely overlapping with SRL's focuses on setting goals, monitoring progress, and adjusting learning outcomes (Andrade, 2019; Andrade & Brookhart, 2016).

However, the role of self-assessment needs to be defined more explicitly and explicated within each phase of self-regulation (Yan, 2020).

It has been argued that self-assessment practices enhance students' reflection on, and monitoring of, their learning progression, and prepare students to internalise self-regulatory skills to promote further long-term learning (Black & Wiliam, 1998; Lam, 2016; Nicol & Macfarlane-Dick, 2006; Panadero, Jonsson, et al., 2016). As the relationship between self-assessment and SRL is not yet robust, the benefits of practicing self-assessment need more research attention (see the review from Brown & Harris, 2013). A later meta-analytic review suggested the positive influence of self-assessment interventions on students' SRL strategies (Panadero, Jonsson, & Botella, 2017), however, the role of SRL measurement needs to be carefully considered as the degree of such positive effects varies with different SRL measurements. For instance, self-assessment seems to have more potential to promote SRL if it is measured with qualitative data such as open questions (Xiao & Yang, 2019).

2.3.1 Self-assessment and learning achievement

The positive effects that self-assessment may have on learning and academic achievement are recognised and there continues to be research interest in the field (see Boud, 1989; Brown & Harris, 2013; Harris & Brown, 2018; for a detailed review). Some researchers have argued that self-assessment is a necessary part of effective learning, (Black et al., 2003; Hobson, 1996), because students need to be able to self-assess their current knowledge, what they know and do not know, to be able to acquire new knowledge. To achieve a radical and sustainable change in students' learning, students need to be encouraged to evaluate their learning (Pintrich, 2002). Likewise, Sadler (1989) claimed that a prerequisite for students to improve their learning is to develop the

capacity to supervise the quality of their learning (González-Betancor et al., 2019), and Pausch and Zaslow (2008) remarked that "the only way any of us can improve is if we develop a real ability to assess ourselves" (p. 112).

It is posited that self-assessment brings a range of benefits to students' learning. Firstly, self-assessment facilitates students to achieve better test performance, a deeper understanding of their work, and a higher task quality (Sluijsmans et al., 1998; Harris & Brown, 2018). Secondly, self-assessment fosters a more responsible and reflective learner compared with peer- or teacher-assessment (Black & Wiliam, 1998). Some researchers have claimed that peer assessment and teacher assessment can serve as a catalyst for students' self-assessment (Dochy et al., 1999; Boud, 1995; Sadler, 1989) by "clarifying where students' misunderstanding lie" (Longhurst & Norton, 1997, p. 320). Thirdly, self-assessment affords students an opportunity to obtain power or control over their learning (Tan, 2004) through learning to assess and assessing to learn, in contrast with learning contexts in which they have no power (Chen, 2008). In doing so, students take the responsibility for their learning and are committed to improving themselves as learners (Crooks, 2011; Dann, 2014; Earl & Katz, 2006; Pintrich, 2002).

Because of the claimed benefits, the emphasis has been placed on the learning function of self-assessment (Bouziane & Zyad, 2018; Sambell et al., 2013), to scaffold students from "dependent novices and consumers of information to becoming agentic individuals able to participate in the social notion of determining quality work" (Tai et al., 2018, p. 478), through fostering their capability in generating formative and evaluative judgements of their work. Whereas self-assessment can be used in a summative manner, maximum benefits can only be realised when used formatively (Andrade, 2019; Yan, 2018). In other words, self-assessment promotes students' learning by deeply engaging

them in evaluating and understanding their own knowledge or skills rather than generating scores for themselves (Birjandi & Siyyari, 2010; Harris & Brown, 2018; Panadero et al., 2019; Yan, 2018).

Based on social cognitive theory (Bandura, 1997), as represented in Figure 2.2 the benefits of positive self-assessment have also been associated with students' higher accomplishment, which is the outcome of students' goals and effort (Ross, Hogaboam-Gray, & Rolheiser, 2002).

In the cycle shown in the diagram, three steps make up self-assessment (shown in the inner box). Firstly, students observe their performance with a particular focus on whether or not it has reached their perceived standards of success. Secondly, students judge how well they have achieved their goals, and then, students finally respond to the results to identify their satisfaction levels based on previous actions. Students' confidence is therefore promoted and leads to more specific goal setting and corresponding efforts, crucial for further achievement (Ross, 2006).

Figure 2. 2. How Self-assessment Contributes to Learning

Note. From Ross et al. (2002, p. 45)

Some studies have examined empirically the effects of teaching students how to self-assess their work in a classroom setting with the results consistently demonstrating a medium effect size gain on students' achievement (e.g. Boud & Falchikov, 1989; Brown & Harris, 2013). In a meta-analytic study, Hattie (2009) investigated more than one hundred instructional and contextual factors concerning their influence on students' educational achievement and reported that self-assessment had the highest effect size on students' outcomes. Yan and Brown (2017) explained the reasons for such strong positive gains, contending that "the benefits of self-assessment may come from active engagement in the learning process, rather than by being veridical or coinciding with reality, because students' reflection and metacognitive monitoring lead to improved learning" (p. 1248). The following sections review studies on the use of self-assessment in the L1, the EFL, and the Chinese writing contexts.

2.3.2 Use of Self-assessment in the Writing Context

Over the last three decades, there has been increasing research investigating the role of self-assessment in writing at different learning levels (Boud, 1989; Lam, 2016; Lee, 2017; Nielsen, 2012; Sadler, 1989), although it is considered to be a research field in its infancy (Hawe & Parr, 2014; Nielsen, 2019). Two key areas were identified in previous studies concerning self-assessment in writing: the accuracy of students' self-grading in relation to that of the teacher's grading, and the effect of using self-assessment approaches on students' overall learning. The current study aims to follow both strands with self-assessment as a mechanism for students to assess, to analyse and to improve their writing performance.

Self-assessment and writing should be closely connected with one another for several reasons. Firstly, as a process itself, writing is thought to be an effective platform to

practice self-assessment because there is a tangible reference when students are analysing and evaluating their written work (Hobson, 1996). Secondly, the act of self-assessment could help students develop a proactive and critical stance when learning writing and judging their writing progress (Lam, 2016). Finally, Lee (2017), informed by Black and Wiliam (1998), who argued that "self-assessment by the student is not an interesting option or luxury; it has to be seen as essential." (p. 55), re-iterated the idea that self-assessment ought to be a fundamental practice integrated into every writing classroom, not be an occasional practice.

A substantial body of research has also discussed how self-assessment could be used in the writing context. For instance, after reviewing the benefits of using self-assessment strategies in writing instruction, Nielsen (2012) highlighted three vital writing self-assessment practices: self-assessment according to particular written prompts; open-ended writing reflection, and oral presentation of writing. Likewise, Lee (2017) pointed out three possible forms of self-assessment in writing, including self-assessment based on writing criteria, self-editing the written language and the organisation, and students' self-inquiry regarding writing achievement and future goals. The employment of these self-assessment practices could help students obtain more specific information about their writing. The most prevalent practice adopted in previous studies, however, seems to be self-assessment using rubrics, which have been reported to influence students' writing quality positively (e.g., Andrade, Wang, Du, & Akawi, 2009; Matsuno, 2009; Wang; 2016).

2.3.2.1 Research in the L1 context

Most studies on self-assessment of writing over the past two decades have been conducted with middle school students and undergraduates from U.S. and UK, followed

by New Zealand and Australia (for reviews, see Riazi, Shi, & Haggerty, 2018; Zheng & Yu, 2019). For instance, positive correlations were identified when comparing university students' and lecturers' grades as an overall rating in which students and lecturers used the same marking guide to assess students' assignments, even though students were not properly trained (Longhurst & Norton, 1997; Sullivan & Hall, 1997). However, students in the above two studies were less accurate in assessing their argument quality and essay structure due to their lack of understanding of the criteria.

Unlike the studies conducted by Longhurst and Norton (1997), and Sullivan and Hall (1997), self-assessment of writing is generally performed with training, and at lower school levels. Taking Ross et al.'s (1999) quasi-experimental study as an example, Grade 4th-6th students ($n = 296$) were employed to investigate the effects of using self-assessment on students' narrative writing. Students in the treatment group were taught how to self-assess their writing against a predetermined rubric (60 mins/day) for eight weeks, whereas the control group was not. It was shown that, after eight weeks, the weak writers in treatment groups become significantly more accurate in self-assessment, although they only slightly outperformed the control groups in narrative writing performance (effect size = 0.18). The small impact might be because the students were not familiar with the rubrics on which they would be graded, as in classes for self-assessment practice, they used another rubric.

Similarly, but over much shorter training periods, the effectiveness of rubric-referenced self-assessment has been illustrated in the studies by Andrade and her colleagues (e.g., Andrade & Boulay, 2003; Andrade & Du, 2007; Andrade, Du, & Wang, 2008; Andrade, Heidi, Du, & Mycek, 2010) as well as other researchers (e.g., Hanrahan & Isaac, 2001) in both qualitative and quantitative studies. From a qualitative perspective, Hanrahan

and Isaacs's (2001) study explored tertiary students' perception of self- and peer-assessment ($n = 233$). The themes they generated from students' answers included both positive and negative responses to self-assessment and peer-assessment on written assignments. However, since peer and self-assessment were integrated, it is difficult to achieve a deeper understanding of students' views on self-assessment only. A later study conducted by Andrade and Du (2007) provided a relatively in-depth understanding of undergraduate students' experiences of using rubric-referenced self-assessment. In focus group interviews, students expressed positivity toward self-assessment reporting that they believed in the benefits self-assessment could bring to their work quality and assignment grade. Students also pointed out the possible teacher (power) influence, that is, mostly the teacher's expectations, on their self-assessment process.

Andrade and her colleagues (Andrade & Boulay, 2003; Andrade et al., 2008; Andrade et al., 2010), using quantitative analyses, reported that U.S. elementary and middle school students' writing quality was improved after scaffolding students with self-assessment processes and using rubrics and model papers to self-assess drafts. However, as treatment times in those studies were short, students only self-assessed very few writing assignments without focusing on the self-assessment process; the effect of such a treatment in a longer term was therefore in doubt. Girls tended to receive higher, but not statistically significantly higher, scores than boys, but as the sample size for girls in most of the studies greatly outnumbered boys, it was difficult to predict the influence of gender on writing scores during self-assessment.

In a recent study, Wanner and Palmer (2018) surveyed students' experiences and understanding on the effectiveness of self- and peer-assessment in geography courses. Similar to what was found in Hanrahan and Isaac's (2001) study, their findings also

indicated that students were mostly favourable to the benefits in terms of quality that practicing self-assessment brought to their writing. However, there were also criticisms from students who considered self-assessment as arduous, time-consuming and frustrating. Furthermore, there was no significant improvement detected in students' learning outcomes, as reflected in students' unchanged assignments' marks after practicing self-assessment during the course.

2.3.2.2 Research in the EFL context

Early research on self-assessment appears to have paid little attention to how it might contribute to EFL writing but, recently, the positive effects of using self-assessment have led researchers to consider its application in the EFL writing context. In particular, there has been an increasing number of studies in Asian and African countries (e.g., Birjandi & Tamjid, 2012; Bouziane & Zyad, 2018; Liu & Brantmeier, 2019). As well as examining the agreement between students' and teachers' grading, there have been studies investigating students' and teachers' perceptions and experiences during self-assessment of writing (e.g., Burner, 2016; Wang, 2016). However, empirical research studies in EFL university level contexts remain underexplored due possibly to the belief that university students are resistant to self-assessment (Harris, 1997; Meihami & Razmjoo, 2016), or the potential tension self-assessment poses in the 'teacher-controlled' learning contexts (Carless, 2011).

Extending previous research (Mok et al., 2006; Ross et al., 1999) that asserted the importance of continual training in the development of accurate self-assessment, Oscarson (2009) and Burner (2016) examined secondary students' and teacher's experiences of integrating self-assessment into everyday writing practice in FL contexts in Sweden and Norway. While their qualitative studies further affirmed the benefits of

using self-assessment in the writing classroom, any benefits from self-assessment may require a longer time for students and teachers to accept. Two explanations were posited: While teachers acknowledged the importance of self-assessment, but they were too conservative to integrate it in their writing class practices; whereas the majority of the students had engaged in self-assessment, they felt there was no benefit in the short term and had not enjoyed the process of self-assessment. Oscarson (2009) also attempted to analyse students' writing linguistically in terms of linguistic accuracy, although the effect of self-assessment on other linguistic indexes of students' writing, such as accuracy was not specifically reported.

A series of studies in the writing domain, conducted in the Asian context, concur with previous studies discussed. These studies investigated the effects of applying self-assessment in various forms and include: self-assessment used alone (Butler & Lee, 2010；Mazloomi & Khabiri, 2018); self-assessment with peer-assessment (Birjandi & Siyyari, 2010), and self-assessment with both peer- and teacher-assessment (Birjandi & Hadidi Tamjid, 2012; Matsuno, 2009). Although some of these studies (Birjandi & Hadidi Tamjid, 2012; Birjandi & Siyyari, 2010) reported considerable benefits from implementing self-assessment in the EFL writing class, on students' writing performance, the mixed findings were not helpful in identifying the effects of implementing self-assessment alone as self-assessment was only part of a multi-pronged intervention. To investigate the potential effect of using self-assessment alone on student' learning, a recent study, Mazloomi and Khabiri (2018) used a quasi-experimental approach with only self-assessment as an intervention to improve students' writing skills and language proficiency. Their findings concurred with those of Mok et al. (2006) who found that students' continuous training and teachers' appropriate feedback during self-assessment are the prerequisites for improvement in writing.

2.3.2.3 Research in the Chinese context

A small number of studies in China have examined empirically the use of self-assessment in the writing domain. For instance, in Hong Kong, Lee and her colleagues have been using assessment for learning approaches (self- and peer-assessment) with writing in their case studies of secondary students and teachers (Lee, 2011a, 2011b; Lee & Coniam, 2013). Their studies not only demonstrated the positive effects of implementing formative assessment with respect to secondary students' attitudes to writing and writing performance, but also indicated the dilemma teachers experienced when faced with traditional and innovative assessment practices. In those studies, self-assessment was implemented only in a limited form, such as self-editing according to the prepared checklists; it was not sufficient to engage students fully in their writing processes and products. It was difficult to identify exactly how self-assessment influenced students' writing performance. The importance of teacher feedback has been illustrated through Lam's (2013) study, with Grade 13 Hong Kong graduates ($n = 31$), which investigated the relationship between self-, peer, and teacher assessment and text revision. The results indicated that, when teachers provide students with focused feedback concerning their self-assessment, students' final drafts were likely to improve. Though the link between self-assessment and text-revision is not straightforward, the role of self-assessment in improving students' texts was confirmed.

Hong Kong is a different context to mainland China as its education system and learning assessment culture are influenced by British traditions (Bai, Chao, & Wang, 2019). In many of the previous studies in L1, in mainland China and Hong Kong contexts, self-assessment was performed with young learners in the classroom, and there are a limited number of studies focused on the university level students (e.g., Liu, 2002; Wang, 2016;

Zheng, Huang, & Chen, 2007). For instance, in what appears to be the first study using self- and peer assessment in the writing class, Liu (2002) investigated the reliability and validity of self- and peer assessment with second-year university students ($n =120$). In Liu's study, English major students were deemed a high English proficiency group and Japanese major students were considered a low English proficiency group. Without any training, all the students self-assessed and peer-assessed one piece of argumentative writing, and their rating was compared with the teachers' rating. Liu indicated that Chinese university students' self-assessment reliability appeared to be affected by students' English proficiency. Self-assessment, therefore, might be more suitable for the high-level group than the low-level group because low-level students tended to overestimate their writing whereas high-levels students seemed to assess themselves objectively. However, the approach Liu used to simply define English-major students as the high-level group and Japanese-major students as the low-level group concerning English proficiency seemed problematic. Moreover, as students from the two groups wrote essays on different topics, it was difficult to demonstrate which group of students' assessment was more accurate.

Similar to the L1 context, in mainland China, some researchers have also examined the effectiveness of rubric-referenced self-assessment in tertiary writing classrooms with non-English major students (e.g., Wang, 2016; Xu, 2019; Zheng, Huang, & Chen, 2007). For instance, over an eight-week period, Zheng, Huang, and Chen (2007) taught Chinese undergraduate students ($n = 189$) self-assessment skills, after which students were required to self-assess three writing tasks using the College English Test writing criteria (from a Chinese National English Test). After eight weeks, they found both students' writing performance and self-assessment accuracy (as compared with that of the raters' rating) improved significantly. Their study affirmed the value of using a rubric to help

in the self-assessment of writing; however, it is not clear whether students from different proficiency levels had made similar gains in their writing after engaging in self-assessment, and whether students' positive gains in writing could sustain over a longer period.

To investigate the long-term effect of using rubric-referenced self-assessment in the writing course, Wang (2016) engaged 80 Chinese undergraduate students in three 45-min training sessions before a further period of 32 weeks practicing both self-assessment and peer-assessment of writing with a rubric adopted from Jacobs, Zingraf, Wormuth, Hartfiel, and Hughey (1981). Findings from interviews and reflective journals not only indicated students' positive views towards using the rubric in self-assessing their writing but also identified factors that may hinder students to realise the full efficacy of the rubric. Although Wang's (2016) study focused mainly on how students used the rubric during self-assessment, it was useful to gauge Chinese students' responses to varied perspectives of self-assessment of writing. Later, adopting a quasi-experimental approach but without any teacher involvement, Xu (2019) further examined whether using rubrics with added annotations for self-assessment of writing could enhance students' writing performance. Findings from Xu's (2019) study were positive, with increased students' writing scores and enhanced students' writing confidence, however, students' interview results suggested the importance of teacher feedback or guidance during self-assessment practices.

More recently, some researchers have focused their attention on the self-assessment of writing with young Chinese learners (ages 12 to 14). For instance, Liu and Brantmeier's study (2019) demonstrated that Chinese young learners could self-assess their English reading/writing skills and knowledge accurately as the correlations between self-

assessment of reading/writing scores and reading comprehension/writing test scores were both found to be significant. Their findings added empirical evidence of young learners' self-assessment trajectories, but they did not explain in any detail about the processes that young learners used to self-assess their writing.

In sum, previous studies employing self-assessment in the writing class have shown its positive effects on Chinese students' writing performance, however, few have implemented systematic self-assessment in the tertiary writing classroom with teacher feedback, nor are English-major students commonly involved.

2.4 Self-efficacy Beliefs, Self-regulated Learning and Social Cognitive Theory

Self-efficacy emerged in the mid-1980s as an aspect of social cognitive theory. It has been acknowledged by researchers during subsequent decades as a strong predictor of academic achievement in foreign language learning and it plays a central role in education and educational psychology (see review, Honicke & Broadbent, 2016; Richardson et al., 2012). Because of the complexities of educational settings, with difficulties in controlling variables and influences, research on self-efficacy beliefs and their effects on behaviour and performance is complex and challenging (Schunk & DiBenedetto, 2016). Self-efficacy is argued to be a salient variable in SRL, playing "a central role in the exercise of personal agency by its strong impact on thought, affect, motivation, and action" (Cliffs, Hall, & Bandura, 1991, p. 248); and, similar to self-assessment, self-efficacy is also integral in all phases of self-regulation in regulating and influencing individuals' thoughts and actions.

In this current study, it is hypothesised that learners' self-efficacy could affect their actions, thoughts, and situation and, in turn, learners can shape or affect their

accomplishments by carefully orchestrating their beliefs, behaviours, and surrounding environments (Bandura, 1997; Kim, Wang, Bong, & Ahn, 2015; Usher & Pajares, 2008). The following sections firstly review empirical studies that have examined self-efficacy beliefs about writing in the L1, EFL and Chinese contexts. As both self-efficacy and self-assessment imply people's judgement about their capabilities, learning, attainment and performance (Bandura, 1986; Boud & Falchikov, 1989; Zimmerman, 2000), the interplay between self-efficacy and self-assessment will also be discussed.

2.4.1 Measuring self-efficacy about writing

In the writing domain, learners' self-efficacy, as an affective factor, is crucial because it predicts individuals' academic achievement through students' capability to understand or to produce varied types of texts with proper grammar usage (Pajares & Johnson, 1994; Pajares & Valiante, 1997; Teng, Sun, & Xu, 2018) and that it can powerfully affect all phases of writing (Boscolo & Hidi, 2007). Self-efficacy also links to students' ability to address emotions that accompany writing and to evaluate their outcomes (Bruning, Dempsey, Kauffman, Mckim, & Zumbrunn, 2013). With a substantial body of research examining how efficacious learners are in specific writing skills and writing tasks (e.g., McCarthy, Meier, Rinderer, 1985; Pajares & Valiante, 1999; Shell, Murphy, & Bruning, 1989), and self-regulatory competence (Zimmerman & Bandura, 1994), recent research has examined writing self-efficacy's multi-dimensional features in both L1 and EFL contexts (e.g., Bruning et al., 2013; Chen & Zhang, 2019; Teng et al., 2018).

When measuring self-efficacy beliefs about writing, previous research mainly used self-report, with questionnaires as the primary tool to assess self-efficacy from a broad perspective, which included writing-related skills and tasks (Chen & Zhang, 2019; Kim, Wang, Ahn, & Bong, 2015; Pajares, 2003; Pajares & Valiante, 2006; Schunk, 2003;

Shell et al., 1989; Woodrow, 2011). Data from such approaches could be problematic as learners' self-efficacy for specific dimensions of writing has not been examined and more importantly, only represents what learners claim they do, and does not provide information about what they actually do. Despite a unitary data collection approach (self-report questionnaires), previous studies have then begun to assess writing self-efficacy in more context-specific ways, for example, linking self-efficacy to students' gender, year level, text revision, writing performance and other affective variables such as motivation and anxiety (e.g., Chen & Zhang, 2019; Villalón, Mateos, & Cuevas, 2015; Woodrow, 2011). The following sections review relevant studies concerning self-efficacy in the writing context, with a focus on the measures utilised to operationalise self-efficacy.

2.4.1.1 Research in the L1 context

Many writing self-efficacy studies involve students in the United States in their samples (e.g., Bruning et al., 2013; McCarthy et al., 1985; Shell et al., 1989; Zimmerman & Bandura, 1994). For instance, research by McCarthy and associates (1985) is one of the earliest empirical studies to investigate explicitly the relationship between college students' writing self-efficacy and performance. They used 19 skills (mainly concerning writing mechanics) that students used in writing expository essays as indices of students' writing quality and measured how confident undergraduate students ($n = 137$) were in demonstrating those skills, using students' ratings in self-report questionnaires. Their study showed that, generally, self-efficacy for writing skills are significant predictors of writing achievement.

Likewise, in another seminal study, Shell et al. (1989) developed a questionnaire to measure undergraduate students' ($n = 153$) writing self-efficacy with component skills

and writing tasks. Regression analysis revealed that writing skill self-efficacy predicted writing achievement, but self-efficacy for the writing task did not. Similarly, at the tertiary level, Pajares and his associates also embarked on measuring writing self-efficacy (Pajares, 2007; Pajares & Johnson, 1994; Pajares, Miller & Johnson, 1999; Pajares & Valiante, 1999) and corroborated previous findings regarding the significant correlation between writing self-efficacy and writing performance.

Unlike the studies discussed above that measured writing self-efficacy for broad writing skills and tasks, Zimmerman and Bandura (1994) situated the evaluation of self-efficacy in self-regulatory theory. They devised the writing self-regulatory efficacy scale to examine first-year college students' ($n = 95$) perceived capability to execute writing activities such as planning and revising. Their results indicated the predictive role of students' self-regulatory efficacy for writing, in students' writing achievement and choices of self-evaluative standards. Also, they reported that if students were not confident in their self-regulatory capability, they were likely to give up on the challenging tasks and so unlikely to implement SRL strategies to achieve higher goals.

Later, using the Writing Self-efficacy Scale (WSES), Pajares and his associates extended the findings of prior studies (e.g., McCarthy et al., 1985; Pajares & Johnson, 1994; Shell et al., 1989; Zimmerman & Bandura, 1994) to students from lower grade levels, such as in elementary school (Pajares et al., 1999) and middle school (Pajares & Valiante, 1999). These two studies indicated that another variable, gender, may not influence writing self-efficacy, but girls did perceive themselves as better writers than their peers. As Bandura (2006) suggested that self-efficacy scales "must be tailored to activity domains and assess the multifaceted ways in which efficacy beliefs operate within the selected activity domain" (p. 310), Pajares (2007) analysed the psychometric properties of the

WSES using exploratory factor analysis with data from Grade 4 to 11 students (n = 1258). His study revealed a two-factor solution of writing self-efficacy in basic grammar skills and complex composition skills, across different levels of schooling.

Since then, following the call to consider self-efficacy in a multi-faceted way, scholars such as Bruning and others (2013) proposed and examined a three-factor (writing ideation, writing conventions, and writing self-regulation) model of writing self-efficacy in two studies with middle school students. Their results corroborated the multi-faceted nature of self-efficacy and the positive correlations between writing self-efficacy (in three dimensions) and writing performance.

2.4.1.2 Research in the EFL and the Chinese context

In comparison with research in L1 contexts, studies of self-efficacy about writing in the EFL and the Chinese context are limited, especially in recent years (Chen & Zhang, 2019; Kim et al., 2015; Li, Liu, & Liu, 2013; Sun & Wang, 2020; Tang & Xu, 2011; Teng et al., 2018; Wang & Kim, 2017; Wang et al., 2014; Woodrow, 2011).

For EFL writers, self-efficacy often plays a more prominent role in their writing process than L1 writers because EFL writing "is strategically, rhetorically and linguistically different in important ways from L1 writing" (Silva, 1993, p. 669). Several recent EFL studies have shown that across different educational levels (e.g., secondary schools and universities), students' self-efficacy about writing has a bearing on students' writing performance (e.g., Kim et al., 2015; Ruegg, 2018; Villalón et al., 2015).

To be specific, using an adapted version of the questionnaire used by Shell et al. (1989), Villalón et al. (2015) found that Spanish high school students (n = 111) displayed no gender differences in writing self-efficacy although girls conceived writing in a more

sophisticated sense than their male counterparts. In contrast, using two questionnaires (English self-efficacy and English SRL), Kim et al. (2015) have provided tentative evidence that among Korean undergraduate students ($n = 167$), female students reported stronger writing self-efficacy than male students (Pajares & Valiante, 1997). Though that results have not been conclusive on the relationship between gender and writing self-efficacy, Kim et al.'s (2015) study revealed that highly efficacious students tended to deploy SRL strategies more frequently in the writing process than those students with a lower level of writing self-efficacy.

Regarding the Chinese context, Tang and Xu's (2011) investigation appears to be the first study to examine Chinese undergraduate students' English writing self-efficacy. Data from non-English major students' self-report questionnaires ($n = 218$), indicated that Chinese students had an intermediate level of self-efficacy for writing; and that students' writing performance was affected by both writing task self-efficacy and writing skill self-efficacy. Woodrow (2011), similarly, explored the relationship of writing self-efficacy and anxiety among Chinese non-English major students ($n = 738$), using structural equation modelling (SEM), and suggested that self-efficacy played a mediatory role between anxiety and writing performance. Whereas Li et al.'s (2013) study identified that self-efficacy levels (from 356 first-year Chinese college students) significantly predicted students' use of writing strategies and writing performance.

Consistent with Bruning et al.'s (2013) study, Teng et al.'s (2018) recent study validated a three-dimensional structure of writing self-efficacy in EFL contexts with data from Chinese non-English major university students ($n = 609$). Their findings showed a small to moderate correlation between the three dimensions of self-efficacy and writing performance and significant correlations between writing self-efficacy and motivational

beliefs (e.g., task value and intrinsic/extrinsic goal orientation). In a later study, Chen and Zhang (2019) proposed and operationalised a two-factor model to investigate Chinese undergraduate students' self-efficacy beliefs regarding high-level revision (altering text content or organisation) and low-level writing revision (revise language problems and mechanics). Their results suggested weak but positive correlations between self-efficacy beliefs and writing test scores, which has contributed to understanding self-efficacy beliefs in revising argumentative writing. Nonetheless, much work remains to be done to provide a more comprehensive view of EFL learners' self-efficacy in other writing genres in the Chinese context.

Extending Kim et al.'s (2015) study, Sun and Wang (2020) examined the relationship among English writing self-efficacy, the use of writing SRL strategies, and writing proficiency with Chinese college students ($n = 319$). Their results were in agreement with those obtained by Tang and Xu (2011), reporting a moderate level of writing self-efficacy and infrequent use of writing SRL strategies for Chinese students. Their results also indicated that both writing self-efficacy and writing SRL strategies were significant precursors of students' writing proficiency. Nevertheless, since the student samples that Sun and Wang (2020) used for data collection were homogeneous regarding their writing proficiency, their results may not be representative of Chinese college students.

Previous studies in L1 contexts have identified the significant effect of self-efficacy on learners' emotion (e.g., anxiety), motivational beliefs and writing performance and they have provided indications that may hold for the Chinese EFL context. However, given the complexity of the writing process, encouraging as the above findings of writing self-efficacy are, the present knowledge as to how to measure EFL writing self-efficacy is still limited. To encourage confident and active EFL learners in the tertiary context, more

studies are needed to explore students' writing self-efficacy from a multi-dimensional and skill-specific perspective (Wang et al., 2014).

2.4.2 Interactions of self-efficacy and self-assessment

The relationship between self-efficacy and self-assessment has been considered by a number of scholars (Bandura, 2007; Kim et al., 2015; Ross et al., 1999; Van Reybroeck et al., 2017). Conceptually, self-assessment has been defined as "assessment of self-efficacy" (McCarthy et al., 1985, p. 465), suggesting the importance of including assessment of an individual's abilities and performance in the conception of self-assessment. Likewise, within the social cognitive theory, self-efficacy indicates "people's judgment of their capabilities to organize and execute courses of action required to attain designated types of performance" (Bandura, 1986, p. 391), and that, to some extent, resonates with the characteristics of self-assessment. Similar to self-assessment, self-efficacy beliefs are rooted in the effect of human agency on one's actions (Bandura, 2006). Because of such close conceptual relationships, Pajares and Valiante (2006) commented that self-efficacy is synonymous with self-assessment so that one cannot develop easily without the other.

Further, it has been argued that the self-assessment training helps students' to internalise their performance as a sense of mastery; such beliefs are deemed to be strongly reinforcing self-efficacy which could lead to similar success in the future (Bandura, 1997; Black & Wiliam, 1998; McMillan & Hearn, 2008). Simply, as Ross (2006, p. 4) has claimed, "asking students to assess their performance, without further training, contributes to higher self-efficacy, greater intrinsic motivation and stronger achievement"; therefore, the higher students' self-efficacy, the greater students' perception of the role of effort, leading to positive self-evaluation (Woodrow, 2011).

The research reviewed in this strand has reported that students' positive self-assessment of their capabilities and progress in acquiring writing skills is vital for sustaining self-efficacy for better learning and performance, because students are aware of their potential to progress (Nielsen, 2019; Schunk, 2003). In the Chinese context, although only a few items pertinent to self-assessment were included in Tang and Xu's (2011) self-efficacy scale, their findings identified the significant impact that self-assessment of EFL writing abilities has on undergraduate students' self-efficacy levels. These results imply that when students endeavour to overcome difficulties through self-assessment of EFL writing, their self-efficacy will be strengthened (Bandura, 1997). Similarly, Alishah and Dolmaci (2013) reported regular self-assessment heightened Turkish first-year college students' self-efficacy levels. The value of self-assessment interventions on improving students' self-efficacy (the average effect size = 0.73) has been further supported by a recent meta-analysis (Panadero et al., 2017). It appears that as self-assessment gives students opportunities to set learning goals and prompts to progress towards achieving them (Andrade, 2010; Panadero & Alonso-Tapia, 2013), students are likely to become more confident (Bandura, 1997). Moreover, self-assessment training, together with students' improved specific task performance, is expected to further increase their confidence as well (Brown & Harris, 2013).

As prior studies have relied mainly on self-report questionnaires, the hypothesis concerning the relationship between self-efficacy and self-assessment needs more empirical evidence, especially in a specific domain, such as, EFL writing. More empirical research therefore is needed to interpret the self-efficacy and self-assessment's relationship.

2.5 Chapter Summary

Initially this chapter, in reviewing how assessment and learning approaches developed from AoL, AfL, AtoL to AaL, described the salient role of alternative assessment forms, such as self-assessment. A comprehensive review of self-assessment as a pedagogical tool was presented with a focus on teachers' role in student self-assessment, variability in reports of the fidelity of student self-assessment and challenges confronted by self-assessment implementation. Even though there are a plethora of studies describing the use of self-assessment as an intervention in the writing classroom, the review of literature has shown that insufficient attention has been given to the EFL setting, particularly in China. The final section introduced the role of self-efficacy beliefs in L1 and Chinese writing contexts and discussed the possible synergies of self-efficacy and self-assessment.

Despite previous researchers' insights and contributions in the field of self-assessment, a number of gaps in the literature have been signalled throughout the review. For example, there is a paucity of research regarding self-assessment of writing in the tertiary classroom, especially in relation to what students actually do, think and feel when they self-assess their writing (Andrade & Du, 2007; Hanrahan & Isaacs, 2001). There is limited research on how student accuracy in rating develops in writing and about the effects of self-assessment on specific linguistic aspects of writing such as complexity, accuracy and fluency. Furthermore, there appears little empirical evidence to inform the assessment of writing self-efficacy beliefs, in particular in the specific domain of self-assessment with tertiary English major students. To address the above research gaps, this study investigates the integration of self-assessment and self-efficacy in the Chinese

tertiary writing context to enhance the understanding of student self-assessment in this specific context.

CHAPTER THREE

THEORETICAL FRAMEWORK

The importance of self-assessment and self-efficacy in advancing students' academic development and of the integration of self-assessment in EFL writing as argued in the research literature were discussed in the previous chapter. The purpose of this chapter is to understand more fully the theoretical rationales of the research design and the variables considered, and to describe the role self-assessment plays in learning, from the perspectives of constructivism, social cognitive theory and formative assessment theory. This chapter starts by introducing how constructivism and self-assessment relate to one another in writing contexts. The subsequent sections describe links among SRL, self-assessment, and self-efficacy in EFL writing framed within the social cognitive theory. The final sections consider self-assessment within constructivism with specific attention on the integration of self-assessment, SRL and formative assessment to enhance the positive effects of self-assessment on learning.

3.1 Constructivism and Self-assessment

Classroom assessment practices are conceptualised within the realm of constructivism, specifically cognitive constructivism and social constructivism (Shepard, 2000). Both strands are pertinent to self-assessment from different perspectives and complement one another to theorise an effective learning environment for students (Powell & Kalina, 2009). Piaget (1952) posited that learners would not internalise learning by simply being given the information; teachers can only mediate students' learning by providing them relevant guidance (Orscarson, 2009), and an individual must actively construct his/her own knowledge through integration and accommodation. This concept emphasises the

active engagement of self in self-assessment in which students gradually understand the procedure and the knowledge underpinning a learning task on their own, rather than being told by their peers or teachers (Dann, 2014). In other words, learning only happens when an individual is able to construct his/her own thoughts via a personal process of experiencing and reasoning through new knowledge (Larsen-Freeman, 2009; Oscarson, 2009; Powell & Kalina, 2009; Sadler, 1989). In this sense, self-assessment practices create a setting for students to understand the criteria they will be evaluated upon, and to reflect on their achievement. Therefore, within this setting, students could construct and reconstruct their skills and capabilities to advance their learning performance.

From a social constructivism perspective, effective learning occurs when students operate within their classrooms as a social environment, access assistance from others and interact with others to shape their own knowledge (Orscarson, 2009; Vygotsky, 1962). In self-assessment, students play a central and active role to process feedback from various external sources to construct new knowledge. They may first need feedback messages from their peers and teachers in understanding and creating criteria, conceptualising and interpreting feedback before generating further self-feedback (Black & Wiliam, 1998; Lee, 2017; Nicol & Macfarlane-dick, 2006).

3.2 Social Cognitive Theory and Self-assessment

Social cognitive theory is a comprehensive construct that informs the "triadic reciprocity in which behaviour, cognitive and other personal factors, and environmental events all operate as interacting determinants of each other" and explicates how people develop and regulate their competencies and actions in learning (Bandura, 1986, p. 18). Within the social cognitive theory, a human has five fundamental capacities, namely, the symbolising capability, forethought capability, vicarious capability, self-regulatory
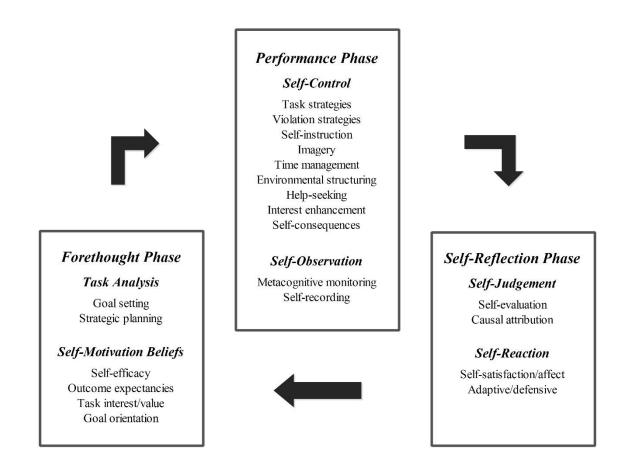
capability and self-reflective capability (Bandura, 1989). Among those competences, self-reflection is a salient and central capability that enables people to self-assess and to further regulate their thinking and judgment. Thus, self-assessment tends to be affected by students' behavioural factors (e.g., mastery experiences), personal variables (e.g., self-efficacy) and environmental factors (e.g., teacher/peer feedback), which in turn are affected by self-assessment (Butler, 2018).

3.2.1 SRL model within social cognitive theory

From a social cognitive perspective, self-regulation is a triadic process in which individuals' personal, behavioural, and environmental components interact to inform how and when to deploy cognitive, affective, and behavioural strategies to achieve favourable learning goals in different contexts (Bandura, 1986; Zimmerman & Schunk, 2011). Zimmerman (2000, p. 14) explained that SRL includes "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals", and self-regulated learners are "metacognitively, motivationally and behaviourally active participants in their own learning processes" (Zimmerman, 2011, p. 49). According to the above definition, the purpose of SRL is to engage students proactively with their learning and motivation management (Zimmerman, 2013).

Among the various models of self-regulation, grounded in social cognitive theory, Figure 3.1, a representative SRL model (Zimmerman & Campillo, 2003, p. 239), demonstrates the working mechanism of the SRL process including three phases: forethought, performance and self-reflection (Zimmerman, 2000). The forethought phase comprises task analysis, which includes goal setting and strategic planning, and motivational beliefs, such as self-efficacy, outcome experiences, task interest/value and goal orientation.

Figure 3. 1.Cyclical Phases of Self-regulation

Performance Phase

Self-Control

Task strategies
Violation strategies
Self-instruction
Imagery
Time management
Environmental structuring
Help-seeking
Interest enhancement
Self-consequences

Self-Observation

Metacognitive monitoring
Self-recording

Forethought Phase

Task Analysis

Goal setting
Strategic planning

Self-Motivation Beliefs

Self-efficacy
Outcome expectancies
Task interest/value
Goal orientation

Self-Reflection Phase

Self-Judgement

Self-evaluation
Causal attribution

Self-Reaction

Self-satisfaction/affect
Adaptive/defensive

Note. From Zimmerman and Campillo (2003, p. 239)

The performance phase features the processes that, through self-control and self-observation, learners employ skills and efforts to optimise their performance. The self-reflection phase involves two main factors: self-judgment and self-reaction. Self-judgment refers to learners' self-evaluation of the effectiveness of their learning or performance against certain criteria or goals, and attributes causality regarding the outcomes. The self-reflection phase provides learners with feedback not only to adjust their learning strategies and goals to adapt to a new cycle of learning but also to influence task analysis and self-motivation beliefs in the forethought phase (Zimmerman, 2013).

In this cyclical process, personal feedback "enables learners to seek assistance when it is needed, expend effort and persist, adjust strategies, and set new goals when present ones are attained" (Zimmerman & Schunk, 2011, p. 1).

3.2.2 Integrating self-assessment with SRL

Nicol and McFarlane-Dick's (2006) study made an initial attempt to explain how assessment practices influence self-regulatory learning processes. More recently, researchers argued for a conceptually intertwined relationship between self-assessment practices and SRL, based on empirical evidence that self-assessment occurs at each SRL phase (e.g., Harris & Brown 2018; Panadero & Alonso-Tapia 2013; Panadero et al., 2018).

Self-assessment is an essential element of self-regulation because they have many commonalities (Andrade, 2010). For instance, both self-assessment and self-regulation prompt students to ponder the quality of their own learning processes and outcomes instead of accepting their teacher as the only source of judgement (Black & Wiliam, 1998; Brown & Harris, 2014). Self-regulation and self-assessment are also believed to be complementary processes for academic improvement (Andrade et al., 2010; Harris & Brown, 2013; Yan, 2020); students' regular involvement in self-assessment encourages academic self-regulation in their own learning, and enhanced self-regulation improves students' self-assessment academic capabilities (Nicol & Macfarlane-Dick, 2006; Panadero et al., 2016; Zimmerman, 2008). For example, using self-report data, Yan (2020) confirmed the fundamental role of self-assessment in different SRL phases by identifying the relationship between different self-assessment actions within and across SRL phases; these included seeking internal feedback, seeking external feedback through monitoring, seeking external feedback through inquiry and self-reflection. He

also reported that, except for seeking internal feedback, other self-assessment actions at the performance phase significantly impact academic achievement, yet, at the self-reflection phase, the academic achievement negatively predicted all self-assessment actions. However, the research design of this study had no control group, which may challenge the credibility of the results.

The relationship between self-assessment and self-regulation can be seen more clearly in the model (Figure 3.1) of self-regulation as shown in section 3.2.1 (Zimmerman & Campillo, 2003). In Figure 3.1, self-assessment plays an indispensable role in the three cyclical phases. For example, in the phase of forethought, self-assessment helps students focus on and revisit their learning goals so that they are more aware of them (Andrade & Brown, 2016). In other phases, Andrade (2010) indicated that two sub-phases are closely connected with explicit self-assessment, namely, self-observation and self-judgement. In self-observation and judgement, students keep track of their own time-to-time performance, and how their performance may influence their further engagement in learning. In this sense, self-assessment underpins and actualises self-regulation by engaging students in a learning cycle consisting of planning, monitoring and judgmental reflection (Clark, 2012). However, given that there are a number of SRL models, it is necessary to identify which SRL model fits and illustrates self-assessment practices more effectively (Panadero et al., 2018).

3.2.3 Self-efficacy in SRL

Resonating with Bandura's (1997) argument that self-efficacy is an individual's judgement about his/her capabilities to deploy actions to accomplish diverse goals, Zimmerman and Schunk (2008, p. 10) described self-efficacy as "the judgement of the ultimate outcomes of one's action". Both definitions emphasise that self-efficacy directs

one's cognitive, motivational, and behavioural skills for goal achievement. In a similar vein, Zimmerman (2008) defined self-regulated learning as "the self-directive processes and self-beliefs that enable learners to transform their mental abilities, such as verbal aptitude, into an academic performance skill" (p. 166). Accentuating the incorporation of processes and beliefs in that definition, Zimmerman (2008) further pointed to the importance of self-efficacy within an SRL cycle. Because, within the social cognitive theory, self-efficacy motivates learners to complete a task with more effort it develops together with learners' cognitive capabilities, and when "self-efficacy judgment increasingly supplants external guidance", the self-regulatory process can be finally realised (Bandura, 1986, p. 414).

Previous research has recognised the indispensable role of self-efficacy in learners' intrinsic motivation and academic achievement (e.g., Pajares, 2003, 2008; Richardson et al., 2012; Schunk & DiBenedetto, 2016; Schunk & Ertmer, 2000). It is contended that, when individuals encounter difficult tasks, high self-efficacy people and low efficacy people react differently in terms of the degree they deploy their resources to cope with the task (Pajares, 2008). For instance, learners with high self-efficacy tend to commit to mastering the challenges with appropriate actions and intrinsic interest (Schunk & Pajares, 2010). In contrast, those with low self-efficacy are inclined to complicate a learning situation, which may lead to unnecessary stress or depression (Pajares & Urdan, 2005).

Within an SRL framework, self-efficacy impacts learners' academic success by continuously regulating and influencing their thoughts, motivation, affect and behaviours. Schunk and Ertmer (2000) indicated the dynamic role of self-efficacy

throughout Zimmerman's three-phase self-regulation model (as shown in Table 3.1). Specifically, Schunk and Ertmer (2000) stated that,

> skilful self-regulators enter learning situations with specific goals and a strong sense of self-efficacy for attaining them. As they work on tasks, they monitor their performance and compare their attainments with their goals to determine progress. Self-perceptions of progress enhance self-efficacy, motivation, and continued use of effective strategies. During the periods of self-reflection, they evaluate their progress and decide whether adaptations in self-regulatory processes are necessary. (p. 634)

Table 3. 1 Self-efficacy in the Phases of Self-regulation

Forethought **(pre-task)**	**Performance** **(during the task)**	**Self-reflection** **(post-task)**
• Self-efficacy • Goals • Outcome expectations • Perceived value	• Self-efficacy • Self-monitoring • Self-perceptions of progress • Strategy use • Motivation	• Self-efficacy • Goals • Self-evaluations • Adaptations of self-regulatory processes

Note. From Schunk and Ertmer (2000, p. 634)

Self-efficacy is an essential construct in the self-regulatory process because only when learners feel efficacious in relation to their capabilities, are they motivated to deploy SRL strategies (Schunk & Usher, 2011).
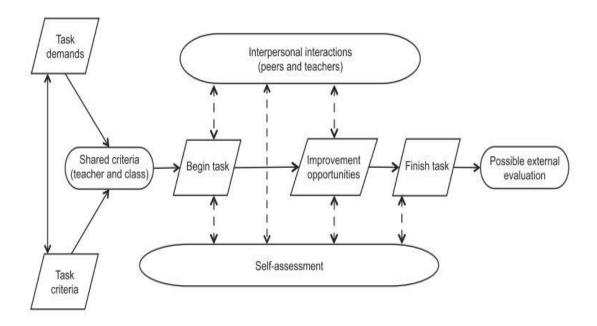
3.3 Formative Assessment and Self-assessment

Formative assessment embodies a multitude of practices, but not all enacted formative practices are with self-assessment. For example, Pryor and Crosouard (2008) distinguished convergent formative assessment from divergent formative assessment;

the convergent formative assessment was "a means by which teachers orchestrated the construction of a lesson text, marking out a correct train of thought for the students to appropriate" (p. 4). In such environments, students receive teacher feedback to ensure their answers are correct, but limited opportunities for self-assessment are allowed. Conversely, the divergent formative assessment practices provide more room for student self-assessment to generate feedback and thoughts in teachers' open-ended questions and knowledge co-construction. As Panadero and Alonso-Tapia (2013) explained, within the formative assessment, "self-assessment is understood as an instructional process used by the teacher as an educational resource" (p. 554).

Within the AfL model, as shown below in Figure 3.2 (Harris & Brown, 2018, p. 27), self-assessment activities fit in the learning cycle and are usually task-related, as assigning tasks to students is considered as an appropriate approach to develop students' self-assessment capabilities (Black & Wiliam, 1998). In Figure 3.2, the solid lines indicate the observable self-assessment, namely, teachers activate students in setting, performing, completing, and evaluating tasks, while the dashed lines mean that how students experience the entire task process and receive peer and teacher feedback is less visible.

Figure 3.2. Self-assessment within Formative Assessment

Note. From Harris and Brown (2018, p. 27)

3.4 Integrating Constructivism, SRL and Formative Assessment in Self-assessment

The above sections have presented three seemingly distinct but overlapping areas of research, which relate to self-assessment in similar ways: constructivism, SRL and formative assessment (Andrade & Brookhart, 2016; Panadero & Alonso-Tapia, 2013). For instance, all theories position students as having key agency during the learning process. However, educators and scholars have tended to view self-assessment from a singular perspective, despite the fact that the effective self-assessment implementation needs to focus on both the internal processes of students and pedagogical strategies (Harris & Brown, 2018). This present study proposes that the three theoretical strands, constructivism, SRL, and formative assessment, work together for a more complete

depiction of self-assessment development to support effective classroom implementation.

Within a constructivism framework, Figure 3.3 demonstrates a simplified integration of formative assessment and SRL around self-assessment when students face a specific learning task. The solid lines refer to teacher-initiated task process and the dashed lines portray learners' self-assessment and self-regulatory activities. In Figure 3.3 (Harris & Brown, 2018, p. 30), the formative assessment bracket at the top indicates students assess themselves formatively for a task while the self-regulated learning bracket at the bottom shows students' active engagement in the task processes. At the beginning of a task, when teachers present the task demands and criteria to students, they discuss the standards and construct background knowledge together. Such discussion first prompts students to self-assess their internal factors, which are relevant to the task, such as goals and competences; then stimulates students to contribute to the criteria and the task as well.

Figure 3. 3. Embedding Self-assessment in SRL and Formative Assessment

Note. From Harris and Brown (2018, p. 30)

As the task proceeds, students self-assess their work quality not only during the task but also at the end of the task against the criteria. Students also monitor their progress and motivation during the course of task achievement. On finishing the task, with self-assessment, students reflect on their work quality to detect possible areas that require further work and deploy appropriate strategies to improve those areas. Then, students feed-forward the knowledge and strategies they constructed from the previous self-assessment into the next task/learning cycle. Interpersonal interactions are also vital when students perform self-assessment given that they need external feedback from teachers and peers to construct further their own self-feedback; a social classroom with healthy interactions among students and teachers is more likely to foster realistic student self-assessment (Harris & Brown, 2018). It can therefore be assumed that the synergy between constructivism, SRL, and formative assessment maximises the effect of self-assessment on students' learning performance.

3.5 Chapter Summary

The chapter opened with a brief explanation of the relationship between constructivism and self-assessment; it then examined how social cognitive theory informed self-assessment from the perspectives of SRL and self-efficacy. The SRL model was introduced in the second part of the chapter where it was shown to be associated with self-assessment. The third part conceptualised self-assessment within the formative assessment theory and this study. Finally, a synergistic model composed of constructivism, SRL and formative assessment, which informs a deeper understanding of self-assessment in EFL writing contexts was introduced. In this study, the perspective of constructivism informs the research design, facilitates the understanding of research questions and data analysis, and addresses the interaction of self-assessment and SRL in

EFL writing. In the specific domain of self-assessment of writing, learners need to construct self-assessment knowledge and capability on their own or co-construct with their teachers to self-regulate not only their relevant behaviours, but also their self-efficacy belief in carrying out those behaviours in the complex self-assessment of writing process.

CHAPTER FOUR

RESEARCH METHODOLOGY

This chapter starts with an overview of the research design, followed by the methodological approach and a description of participants and instruments involved in different stages. The ensuing section offers a brief articulation of the ethical considerations. Then, the following sections provide a detailed explanation of the instruments, participants, data collecting, and analysing procedures involved in each stage of the study.

4.1 Overview of the Research Design

The present study applied quantitative and qualitative methods concurrently to gain an in-depth and focused understanding of the effects of self-assessment and peer-assessment on EFL students' self-efficacy for self-assessment of writing and their writing performance. The study embedded self-assessment and peer-assessment based writing instructions in the traditional writing course syllabus of Chinese tertiary EFL writing classrooms over a time span of four months. It was expected that data from this study would afford rich information about integrating self-assessment into regular English writing classrooms in the Chinese context.

Students' views and experiences in the peer-assessment group were not taken into consideration as other assessment forms are not the focus of this study. Therefore, only views of students who received the self-assessment intervention are presented in the qualitative findings because, I was mindful that, in this study, the original intention was to investigate the effects of self-assessment on EFL students' self-efficacy and writing

improvement and to compare this with a group receiving normal teacher assessment. When I embarked on data collection, however, the research site made an unexpected decision to adopt peer-assessment as a teaching pedagogy in the group that was to form the comparison, rather than continuing with normal teaching, namely, teacher assessment for the upcoming semester, as agreed previously.

The preparatory stage was designed to validate the researcher-developed questionnaire measuring the underlying factors of self-efficacy for self-assessment of writing (details of the validation are presented in Section 4.4) in a broader Chinese EFL setting. Results of the preparatory stage provided guidance for the main study, involving in-depth exploration of self-efficacy for self-assessment of writing. The main study involved two parts. Part one implemented self-assessment-based writing interventions in two EFL writing classes (the intervention groups), and the peer assessment condition (two EFL writing classes) provided a comparison for the self-assessment condition. Data were collected from self-reported questionnaires, pre-and post-writing tests, student self-assessment tasks, and class observations (details for this part of the study are provided in Section 4.4). Part Two was conducted in the self-assessment intervention classes, to supplement and triangulate the findings of Part One with regard to self-efficacy for self-assessment of writing, and to provide additional information on students' experiences in self-assessment of writing. The aim was to achieve a deeper understanding of students' self-efficacy, writing development and their adaptation to self-assessment during the intervention period. To realise the objectives of Part Two, semi-structured interviews and learning journals were used for data collection (see Section 4.5 for more details). Table 4.1 provides a detailed overview of all research procedures.

Table 4. 1 An Overview of the Research Design

Phase	Research Aims	Instruments	Participants
Preparatory Stage	• Instrument development and validation	• Self-efficacy for Self-assessment of EFL Writing Questionnaire (SESAEFLWQ)	• English major students Year 1-Year 4 (N = 556)
Main Study Part One	• Investigating EFL students' current situations in self-efficacy for self-assessment of writing, and writing performance • Implementing self-assessment-based writing intervention. • Examining students' rating accuracy as compared with the raters' assessments. • Exploring the relationship between students' self-efficacy for self-assessment of writing and writing performance over time.	• Pre-and post-Questionnaires (SESAEFLWQ) • Pre-and post-timed writing tests • Pre-and post-writing self-assessment tasks (self-assessment sheet and self-assessment writing rubrics for EFL writers) • Self-assessment-based writing intervention (four months) • Classroom observation	• English major students Year 2 Intervention group (Self-assessment; N = 51) Comparison group (Peer-assessment; N = 41) • English major lecturers Year 2 (N = 2)
Main Study Part Two	• Providing in-depth information about students' change in self-efficacy for self-assessment of writing, writing performance and their adaptation to the intervention. • Comparing raters' and students' writing comments • Identifying factors that influencing students' self-assessment of writing practices	• Semi-structured interviews (pre- and post) • Writing journals (three time points: 4th, 10th and 16th week)	• English major students Year 2 Intervention group (N = 6)

4.1.1 Research paradigm

In applying both quantitative and qualitative approaches in this study, I adopted the pragmatic worldview as my philosophical justification. Pragmatism draws from "actions, situations and consequences rather than antecedent conditions" (Creswell & Creswell, 2018, p10); it implies the use of all possible applications to address the potential problems, employing a range of methods to collect both quantitative and qualitative data to present the best understanding of the research problem (Creswell & Creswell, 2018). A recent review study showed that quantitative methods predominate over other research approaches in the realm of writing assessment (Zheng & Yu, 2019) because utilising quantitative methods is more likely to identify any association and trend between the variables. Although a quantitative approach may not include the subjectivity of people's lives in a certain context (Dörnyei, 2007), this weakness of quantitative approaches can be improved with the inclusion of qualitative methods. Qualitative methods stress the interpretation of individuals' contextual experiences, and therefore are advantageous to "develop a detailed view of the meaning of a phenomenon or concept for individuals" (Creswell & Creswell, 2018, p. 53). By applying quantitative and qualitative methods, I was able to "draw on the strengths and minimise the weakness of both types of research" (Connelly, 2009, p. 31).

As the focal point of this research was the effects of developing self-assessment skills on EFL students' self-efficacy and writing improvement, the pragmatic worldview was best suited in this context because such a view favours the use of a range of supplementary methods to address better the diverse research questions raised earlier in the Introduction.

To investigate the effect of using self-assessment on self-efficacy and writing performance among EFL learners, an observational and quasi-experimental approach was considered most appropriate at the initial stage. Following Loewen and Plonsky (2016), an observational and quasi-experimental design provides rich information for a particular context and makes the comparison between two groups possible. In light of that, this study employed four intact classes from the focal university, to form two participating groups, which were randomly assigned as the intervention group and the comparison group.

To triangulate further the data achieved at Part One, I adopted a case study approach at Part Two. According to Yin (2018), a case study design is appropriate not only when the major research questions concern either "how" or "why" questions, but also when the researcher possesses limited or no control over individual behaviours. There are two reasons why I adopted the case study approach: first, the key research question for Part Two was a "how" question in nature; second, I was not able to control students' lived experiences in self-assessment of writing due to the dynamicity and complexity of students' varied interactions with internal and external factors (Andrade & Du, 2007).

4.1.2 Participants

In total, using convenience sampling (Creswell & Creswell, 2018), this study recruited 668 voluntary student participants and two lecturer participants from 15 medium to large scale universities in three rounds for different parts of this study. These universities, in the northern, middle and southern areas of China, had student enrolments ranging from 20,000 to 60,000. The sample comprised 121 (18%) male and 549 (82%) female participants, with an age range of 17-23 years (M =19.8, SD =1.539). Participants all hold Chinese nationality and speak Mandarin as their first language; at the time of this

study they had an average of 9.82 (*SD* = 1.53) years of formal English learning. Table 4.2 presents information on the participants in different stages. More information about participants involved in every part is presented in later sections.

The sample was all Chinese English-major students recruited through convenience sampling. Although using convenience sampling may not yield a representative population, it was a practical, time-saving and advantageous method to gain willing participants for this research (Dörnyei, 2007; Merriam & Tisdell, 2016). Further, by involving participants from different areas of China, the sample was a reasonably representative of Chinese undergraduate English-major student populations, who are usually of a similar educational background and age.

Table 4. 2 An Overview of Participant Information for the Whole Study

Phase		Research Objectives	Total *N*	Major	Grade	Gender	*M age (SD)*
Preparatory Stage		Instrument development and validation	N = 556	English	Year 1 (18%) Year 2 (34.5%) Year 3 (27.2%) Year 4 (20.3%)	Males (18%) Females (82%)	*M* = 19.9 *SD* = 1.54
		Piloting instruments	N = 10	English	Year 2 (50%) Year 3 (50%)	Males (30%) Females (70%)	*M* = 19.6 *SD* = 1.54
Main Study	Pilot study	Trying out the instruments	N = 10	English	Year Two	Males (30%) Females (70%)	*M* = 19.5 *SD* = 1.12
	Part One	Self-assessment writing intervention	Intervention group (Self-assessment; N = 51) Comparison group (Peer-assessment; N = 41)	English	Year Two	Males (16%) Females (84%)	*M* = 19.2 *SD* = 1.32
			English lecturer (N = 2)			Females	*M*= 38.5
	Part Two	Multiple-case studies	Intervention group (N = 6)	English	Year Two	Males (N = 2) Females (N = 4)	*M* = 19.5 *SD* = 1.27

4.1.3 Instruments

In this study, data were collected from self-report questionnaires, timed writing tasks, student self-assessment tasks, semi-structured interviews, learning journals, and class observations. These instruments were used to evaluate students' self-efficacy for self-assessment of writing, their writing performance, rating accuracy, as well as their response to the intervention process. Further details about the instruments can be found in the relevant section of this methodology chapter.

4.1.3.1 Self-report questionnaires

The self-report questionnaire is an efficient instrument because of its cost-effectiveness and versatility when gathering factual, behavioural and attitudinal information on a large scale (Dörnyei & Taguchi, 2010). With data garnered through questionnaires, it is possible to "provide important insight into general tendencies in a particular population" (Petrić & Czárl, 2003, p. 209). Self-report questionnaires are not exempt from limitations as respondents' careless responses or misunderstandings may lead to poor quality results (Dörnyei, 2007). The inclusion of interviews, however, can complement the questionnaire data by providing in-depth views on the topic (Loewen & Plonsky, 2016). Self-report questionnaires, therefore, are appropriate to examine students' self-efficacy for self-assessment of writing in the EFL context.

4.1.3.2 Timed writing tests

The College English Curriculum Requirements (Ministry of Education, 2018) emphasises students' writing abilities for students to be able to communicate effectively with other parties in their future work. Essay writing is also an essential and effective approach to evaluate students' critical thinking and linguistic abilities when

communicating in written English; it is also used in Chinese national English assessments carrying a high weighting (Liu, 2010). In this study, two cause and effect essays, with similar set topics, were used to examine writing performance and linguistic development of the students from the intervention group and the comparison group. A cause-effect essay was selected as is widely used not only in the Chinese context but also worldwide in EFL classrooms and it constituted a large percentage of the essay types included during the intervention period.

4.1.3.3 Self-assessment writing rubric

Rubric-guided self-assessment is a widely endorsed form of self-assessment owing to its potential in helping students to rate their work, and to generate complex judgements about their work along a continuum (Andrade, 2010; Andrade & Valtcheva, 2009; Brown & Harris, 2013; Harris & Brown, 2018). Defined as "a coherent set of criteria for students' work that includes descriptions of levels of performance quality on the criteria" (Brookhart, 2013, p. 4), rubrics help students to conceptualise a quality piece of writing (González-Betancor et al., 2019; Sadler, 2009). Moreover, being evaluative and instructional in nature (Wang, 2016), formative rubrics have the potential to promote students' learning and enhance students' scoring reliability when self-assessing (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Panadero, Jonsson and Strijbos (2016) argued for explicit rubrics as part of effective assessment practices, for students to be able to make just-in-time adjustments to their learning goals and plans. In this study, students from both the intervention and the comparison groups not only utilised an adapted rubric in the pre- and post- writing tests, but also referred to the rubric in the writing class when self or/peer-assessing their writing tasks (see section 4.4.1.3 for details of the rubric).

4.1.3.4 Semi-structured interviews

In the form of "verbal questionnaires", semi-structured interviews are "a series of questions designed to elicit specific answers from respondents" (Fraenkel, Wallen, & Hyun, 2015, p. 451) to glean a deep understanding of people's views and experiences. Without limiting the depth and breadth of participants' responses, semi-structured interviews not only allow researchers to focus on the interview protocol, but also help researchers to cross-check data elicited with other instruments, such as close-ended questionnaires (Dörnyei, 2007; Merriam & Tisdell, 2016). In this study, semi-structured interviews were conducted with six students from the intervention group (self-assessment) concerning their self-efficacy in relation to self-assessment of writing, writing performance, self-assessment rubric, and their adaptation to the self-assessment intervention. The data from the interviews helped to triangulate and supplement the quantitative data.

4.1.3.5 Learning journals

Learning journals, as Boud (2001) contended, function as "a device for working with events and experiences in order to extract meaning from them" (p. 9), through which individuals can enhance what they do and how they perform certain tasks. Keeping journals is also deemed as an ideal type of self-assessment for students to generate reflective self-comments of their recent performance (Brown & Abeywickrama, 2019; Wang, 2016). In this study, six students from the intervention group (self-assessment), who participated in the multiple-case study, were required to keep learning journals at three time points, the 4th, 10th and 16th week of the intervention. Those three time points were chosen for two reasons. Firstly, they were at the beginning, the middle and the end of the semester, and thus would reveal participants learning and attitudinal trajectories,

and possible changes in the course of the intervention. Secondly, although a writing journal is a retrospective account, which allows the researcher to access learners' internal activities with little intrusion, completing a journal is demanding for the respondents (Dörnyei, 2007), therefore, it was not feasible to ask for more than three instances. I collected the completed journals at each point in time, made a copy of those journals, and then returned the originals to respondents.

4.1.3.6 Classroom observation

Classroom observation not only enables authentic data to be gathered from the targeted population (Creswell & Creswell, 2018), but also captures the complex nature of the classroom (Wright, 2006). As a non-participant observer, I conducted monthly classroom observations (eight lessons in total: four lessons with the intervention group and four lessons with the comparison group) to track, to understand and to evaluate, the classroom practices and students' behaviours during self-assessment. Classroom observations are useful not only in eliciting first-hand information about the intervention implementation, actual teaching practices and students' engagement but also in triangulating the self-reported data achieved from questionnaires and interviews (Cowie, 2009; Merriam & Tisdell, 2016). An aim of the observations in this study was also to ensure that the comparison group students did not engage in self-assessment in their writing class.

4.1.3.7 Raters' and students' comments on writing strengths and weaknesses

As there is limited knowledge of how students learn to assess their own writing, compared with external assessment, the effectiveness of self-assessment should not be contingent on the conformity between students' and teachers' grading. The focus should

be on how EFL students have developed during the assessment process in their ability to identify their strengths and weaknesses (Chen, 2008; Mazloomi & Khabiri, 2018; Orsmond et al., 1997). In addition to investigating the statistical correlation between students' self-grading and raters' assessment, this study examined students' development (from the multiple-case study) in rating accuracy by comparing students' self-comments and raters' comments regarding their strengths and weaknesses.

4.2 Ethical Considerations

This study was conducted after the University of Auckland Human Participants Ethics Committee granted the approval on 7th August 2017 for a period of three years until 7th August 2020 (Reference Number 019574). The researcher does not have any prior relationships with the universities and participants in the study. To address the possible ethical issues effectively, the considerations discussed below were followed.

4.2.1 Informed consent

Prior to completing the online questionnaire voluntarily students were presented with a Participant Information Sheet (PIS), prepared according to the guidelines provided by the University of Auckland Human Participants Ethics Committee, which fully described the research purposes, procedures, and the usage of data. After ensuring the participants understood the research project, they were invited to sign a Consent Form (CF) showing that they agreed to complete the online survey anonymously with genuine answers; participants also indicated whether they wished to receive a copy of the findings when this study was completed.

The researcher sent the PIS and CF to the Dean via email, and with a request for the signed CF to be returned before the study commenced. To avoid coercion in the

recruitment of participants, the research information, in the PIS, was forwarded to the departmental secretary to advertise the research project among the English lecturers and their students to recruit participants. Lecturers and their students interested in the research obtained a PIS and a CF from the secretary and returned the CF with their signatures to a drop-box in the secretary's office. Then the researcher collected participants' CFs from the office. Appendix A contains examples of the PIS and CF for this study.

4.2.2 Participant's right to withdraw

Participants were informed that this research project was a part of the researcher's Ph.D. study and they had the right not to answer any specific questions in the questionnaires or interviews and they were entitled to turn off the recorder during the interview at any time. Also, they had the right to withdraw themselves or any data provided by them from the research at any time before 1st February 2018 without giving a reason. An assurance from the Deans at each university was obtained stating that neither participation nor non-participation in or withdrawal from this study would affect the participants' normal courses, grade, career or relationship with the faculty or the university.

4.2.3 Confidentiality, anonymity, and data retention

To ensure no information could be attributed to an individual or institution, pseudonyms or numeric indexes were used for collecting data, analysing data, and reporting findings. Any materials (records, transcripts, etc.) collected for this study that may expose the identity of the universities and the participants were kept confidential and only known to the researcher during the research period with no disclosure to any third party. Any data collected for this study, used for conference presentations or academic publications,

will use pseudonyms to protect participants' identity. The researcher ensured that any materials collected would not be used to evaluate the individual lecturer's teaching or the student's learning. Hard copies of the research data will be kept for six years in a locked cabinet at the University of Auckland, before being shredded; electronic data would be stored on the researcher's password-protected computer, and permanently deleted after six years.

4.3 Preparatory Stage—Instrument Development and Validation

As no existing questionnaire was available that directly assesses EFL learners' self-efficacy for self-assessment of writing, it was necessary to modify a current instrument or to devise a new instrument for the evaluation of the targeted factors before putting other efforts into it (Wang, Kim, Bai, & Hu, 2014). Since using credible and valid instruments are a precondition of a successful research (Muijs, 2011), a self-reported questionnaire, *Self-efficacy for Self-assessment of EFL Writing* (SESAEFLW) with 25 items was first developed, which was reduced to 20 items after validation. This survey was designed to understand better the underlying factors of self-efficacy for EFL writers in a Chinese context. The following sections describe the instrument development and validating process.

4.3.1 Instrument: Self-efficacy for self-assessment of EFL writing (SESAEFLW) questionnaire

4.3.1.1 Instrument development rationale

Measures of self-efficacy should "reflect an understanding of both the domain under investigation and its different features, as well as the types of capabilities the domain requires and the range of situations in which these capabilities might be applied" (Pajares

& Valiante, 2006, p. 162). Similarly, Bandura (2006) argued that "scales of perceived self-efficacy must be tailored to the particular activity domain of functioning that is the object of interest" (p. 308). However, investigation of self-efficacy in the domain of self-assessment is still rare and the development of a thorough understanding of self-assessment is hindered by the lack of valid instruments for assessing self-assessment practices (Yan, 2020). To date, no models or measures have been developed to elicit information on EFL writers' self-efficacy on self-assessment. The absence of a self-assessment focused self-efficacy scale led me to propose a multi-factor model of self-efficacy that reflects learners' confidence in self-assessing their writing skills and writing performance that could be used in this study.

SESAEFLW consists of two parts: Firstly, participants' demographic information, including students' age, grades and years of English language learning; secondly, students' views on their self-efficacy for self-assessment of EFL writing. Drawing on Bruning et al. (2013), the questionnaire proposed a three-factor model of self-efficacy for self-assessment of writing, which included three key dimensions, learners' linguistic abilities, psychological needs and behaviour demands in the self-assessment of the writing process. The linguistic dimension and the psychological dimensions each comprised five statements; and the dimension of behavioural demands comprised ten statements. The number of items in all three dimensions and the sample size meet the recommended value for factor analyses (Costello & Osborne, 2005). The time for questionnaire completion was roughly 20 minutes (see Appendix B for details of the questionnaire).

4.3.1.2 Item generation

To develop a valid and reliable construct that assesses students' self-efficacy for self-assessment of writing, the item generation process followed Dörnyei and Taguchi's (2010) standard guidelines for questionnaires and considered previous devised self-efficacy instruments (Bruning et al., 2013; Shell et al., 1989; Woodrow, 2011). Items utilised in previous studies were carefully rephrased for self-assessment in the EFL writing context and worded as "I can assess" statements to reflect self-assessments in different situations.

After generating an initial list of 25 items pertaining to self-efficacy for self-assessment of writing, ten EFL students were invited to review the list's clarity and readability. The finalised instrument contained 25 items that were sequenced logically based on the three dimensions. The instrument used a five-point Likert type scale from one (*not at all sure I can do*) to five (*completely sure I can do*), so that the multi-item rating scales provided a "comprehensive content coverage" (Dörnyei & Taguchi, 2010, p. 57). The questionnaire was then scrutinised statistically through factor analyses.

4.3.2 Participants

In the preparatory stage, in total, 566 English-major students from 14 medium to large-scaled universities across China responded and agreed to complete the online survey. Among those participants, ten students joined in the piloting session to improve the clarity and readability of the questionnaire, and 556 students completed the questionnaire online. The sample was obtained through convenience and snowball sampling. Among these participating students, 105 were males and 461 were females. Three male and seven female students in the piloting session were in Year Two and Year

Three (M $_{age}$= 19.6, $SD = 1.54$). The rest of the respondents were in Year One to Year Four (18% freshmen, 34.5% sophomores, 27.2% juniors, 20.3% seniors), with an age range from 17 to 23 years ($M = 19.82$, $SD = 1.54$).

4.3.3 Procedures

In the preparatory stage, ten students first verified the readability and comprehension of the survey's items; 556 English-major students from 14 Chinese universities were then invited to complete the *Self-efficacy for self-assessment of EFL writing* (SESAEFLW) questionnaire through the link of Wenjuanxing, a Chinese survey tool similar to Survey Monkey as widely used in the west. The questionnaire was completed over the period of 08/2017-09/2017 in two rounds, and generated 250 and 306 responses, respectively. It was written in simple English without translation since the respondents were all English major students, able to process the survey items competently in English. The participants were assured that there were no right or wrong answers, as the purpose of the survey was to appraise students' reported self-efficacy for self-assessment of writing. Respondents spent approximately 15 minutes on average for survey completion, as reported in the survey tool.

4.3.4 Data analysis

4.3.4.1 Preparation

Data collected from the Self-efficacy for self-assessment of EFL writing (SESAEFLW) questionnaire were first exported automatically into Microsoft Excel 2013 for screening and cleaning. Following Field's (2013) suggestion, students' mischievous, inaccurate, same or extreme responses were identified with manual inspection and removed from the data pool. The statistical software IBM SPSS Statistics 25 and IBM SPSS Amos 25

were then employed for the subsequent data check, including missing responses, outliers, normality and homogeneity check, and computation using multivariate analyses (exploratory factor analysis, confirmatory factor analysis, and structural equation modelling). Measures of normality, skewness and kurtosis were selected with standardised skewness (from 0 to +/– 3.0) and kurtosis values (from 0 to +/– 8.0), considered as normally distributed (Kline, 2016). To meet the assumption of homogeneous variances for multivariate analyses, the Levene's test was adopted to examine the homogeneity of variance (Raykov & Marcoulides, 2008); all quantitative primary data demonstrated normality and homogeneity of variance.

4.3.4.2 Factor analysis

Factor analysis is effective to cut "a large number of variables to a smaller set of underlying factors that summarize the essential information contained in the variables" (Coakes, Steed, & Price, 2008, p. 128). Hence, after data cleaning, two data sets with 245 and 300 participants, which were collected in two rounds, were subjected to exploratory factor analyses (EFAs) first and then confirmatory factor analyses (CFAs) for instrument validation. EFAs with 245 participants, examined the underlying structure of the newly-developed questionnaire, *Self-efficacy for self-assessment of EFL writing,* which had not been shown in previous literature. Then, a series of CFAs using structural equation modelling (SEM) through IBM SPSS Amos 25 was performed to test if the factorial structure is plausible or consistent with the rest of the 300 participants.

Exploratory factor analyses (EFAs)

The exploratory factor analyses followed procedural guidelines suggested by Pallant (2016). The Kaiser-Meyer-Olkin's (KMO) test and Bartlett's test of sphericity were first

adopted to verify the sampling adequacy. Acceptable KMO values are between .50 and 1 and, the higher the values are, the greater the common variance is (Ferguson & Cox, 1993). The maximum likelihood (ML) estimation via oblique rotation and the parallel analysis was then performed to scrutinise the model fit (Fabrigar, Wegener, MacCallum, & Strahan., 1999) while maintaining the number of components (O'Connor, 2000). Items with low loadings (< .30) and complex loadings were excluded from further analyses. Cronbach's alpha reliability estimates were checked for the three factors with the threshold value (> .70), which indicates satisfactory internal consistency reliability of the instrument (DeVellis, 2012).

Confirmatory factor analyses (CFAs)

Given the sensitivity of structural equation modelling (SEM) to multivariate normality, Mahalanobis distance was examined for the removal of further outliers (Kline, 2016), using SPSS Statistics 25. As a conventional but powerful approach, SEM concurrently measures the degree and directionalities of factor correlations among all items in the data matrix. To evaluate the goodness-of-fit of the proposed model, this study interpreted the CFA data with several omnibus fit indices, including the absolute fit indices and incremental fit indices (Hu & Bentler, 1999). By doing so, possible limitations of one index could be complemented by another index. Specifically, this study reported three absolute fit indices, which indicate the degree of *a priori* model fit, or data reproduction (as shown in the following).

(a) the ratio of chi-square divided by the degree of freedom (*df*).

(b) the Root Mean Square Error of Approximation (RMSEA).

(c) the standardized root mean square residual (SRMR).

Moreover, two incremental fit indices were reported, namely, the comparative fit index (CFI) and the Tucker-Lewis index (TLI) (Tucker & Lewis, 1973). To determine the fit of the measurement and structural equation models, I consulted a number of previous reviews on the acceptable levels of goodness-of-fit indices (e.g., Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999; Schreiber, Nora, Stage, Barlow, & King, 2006). Specifically, the following criteria were used as acceptable levels of fit in the results section: a non-significant chi-square value; $\chi2/df < 3.0$; TLI and CFI $>.90$; SRMR $< .05$; and RMSEA $< .08$. The above test results will be presented in Chapter Five.

4.4 Part One – Self-assessment-based Writing Intervention

A number of studies have affirmed the positive role of self-assessment in a classroom setting (e.g., Andrade & Brown, 2016; Kearney, 2013), with some intervention studies assessing the "effectiveness of teaching methods, curriculum models, classroom arrangements, and other efforts to influence the characteristics of individuals or groups" (Fraenkel et al., 2015, p. 16). Other studies have argued that involving alternative assessment tasks in the target course is an effective approach to improving students' learning (Sluijsmans et al., 1998).

However, there has been limited attention given to how self-assessment functions in EFL settings, particularly in the writing context. Part One of this study, therefore, implemented and evaluated, in an EFL writing context, a self-assessment-based intervention, with a comparison group employing peer assessment that had already been implemented as part of classroom practice. A quasi-experimental research design, comprising two contrasting conditions: the intervention group (self-assessment), and a comparison group (peer-assessment) was used with quantitative data gathered from the

pre- and post-questionnaires, writing tests, and students' self-assessment tasks. The following questions were posed:

1) What effect does the use of self-assessment have on students:

   a. self-efficacy for self-assessment of EFL writing?

   b. writing performance as well as their linguistic improvements in terms of complexity, accuracy, and fluency?

   c. rating accuracy in self-assessment of writing as compared to the raters' assessment?

2) How does self-reported efficacy for self-assessment of writing link with students' writing performance?

4.4.1 Instruments

4.4.1.1 Self-efficacy for self-assessment of EFL writing questionnaire

The *Self-efficacy for self-assessment of EFL writing questionnaire* was administered to all the participants in English. The survey comprised two components: demographic information and the finalised questionnaire (validated in the preparatory stage with satisfactory psychometric properties). With 20 items, the finalised questionnaire examined students' self-perceived efficacy for self-assessment of writing in the EFL context from two perspectives: writing skills and writing tasks self-assessment (Appendix B includes the content of *Self-efficacy for self-assessment of writing questionnaire*).

4.4.1.2 Pre-and post-writing tests

To explore students' writing performance before and after the intervention, students were asked to complete two given-topic cause and effect essays, writing at least 200 words in response to the prompt (e.g., topic information outline) within 45 minutes. The cause and effect type of essay is considered as a common writing task for Chinese undergraduates (Ding & Zhao, 2019). To ensure comparability for all the participants, writing topics, matched in difficulty level and relevance to students' daily lives, which stemmed from previous papers in the Chinese National Test for English Majors, Band 4 (TEM 4) were selected. The topics were: a). What are the reasons for people shopping online and what effects does it have on people's lives and the society? (Pre-test) b). What are the reasons for people shopping in physical stores and what effects does it have on people's lives and the society? (Post-test) Appendix C includes the writing topics. Participants were required to complete the writing tasks independently in class with a pen and paper without any external help; they were at liberty to add, to delete, or to cross out their content, rather than to re-write or erase.

4.4.1.3 Self-assessment writing rubric

In the pre-and post-writing tasks, students from both intervention and comparison groups were provided with a four-dimension self-assessment writing rubric with descriptors for each dimension, which was adapted from previous analytical scoring schemes (Glasswell, Parr & Aikman, 2001; IELTS, 2015; Jacobs et al., 1981). Further, to ensure the rubric was appropriate to the research context, the adapted rubric took account of factors that Chinese EFL students perceived as affecting the rubric's effectiveness during their self-assessment of writing (Wang, 2016), for example, the relatively narrow 'three-point' scoring range was replaced by a five-point one.

Specifically, the rubric comprised descriptors in four dimensions: task achievement, coherence and cohesion, language resources, and mechanics. Each dimension represented a score from one to five, in which a score of three represented writers at a basic level, score four proficient writers, and score five advanced writers. Students were required to self-assess their cause and effect essays according to the above dimensions, generating an individual section score and an overall score according to the provided rubric. Students were also asked to comment on strengths and weaknesses in their writing. To do this, students first indicated a dimension that was a strength and one that was a weakness in their writing, and then explained the reasons behind their choices. Students completed the self-assessment task within 20 minutes. Raters used the same rubric to grade students' pre-and post-writing texts and identifying as well as commenting on students' strengths and weaknesses (see Appendix D for details of the self-assessment writing rubric).

4.4.1.4 Classroom observation

As a non-participant observer, I conducted monthly classroom observations with both groups (four times in total for each group) for two reasons. Firstly, I aimed to ensure the valid implementation of the intervention by checking that the two groups focused only on their assigned assessment form, namely, self-assessment for the intervention group and peer-assessment for the comparison. Secondly, as observation provides "the conscious noticing and detailed examination of participants' behavior in a naturalistic setting" (Cowie, 2009, p. 166), I focused on students' engagement during the intervention, particularly for the six students who were the participants in Part Two. When conducting classroom observation, I kept detailed field notes of teaching practices using a classroom observation sheet (see Appendix E for details), on which I recorded

the use of self-assessment, instructional processes, teacher feedback, classroom activities and students' engagement. The observation sheet and field notes were used as a reference to verify and to complement data collected from other sources.

4.4.2 Participants

The main study consisted of 94 participants in total: 92 student participants and two lecturer participants. Details of the participants are described in the following.

4.4.2.1 Student participants

The 92 students, 15 male students (16%), and 77 female students (84%), from four intact classes were recruited voluntarily from a medium-scaled (around 20,000 students) Chinese university. At the time of this research, these English major students, with comparable years of English learning experience (M years = 9.18, SD = 1.24) and of similar age (M = 19.42, SD = 1.15), had just embarked on their second year of study. They were all enrolled in a compulsory English writing course, which focused on linguistic knowledge and various types of English writing.

The participants reported that they had no overseas learning experience and they had never received any instruction concerning self-assessment of writing in their previous education. As homogeneity among student participants was mostly ensured prior to the research, the potential risk of the convenience sampling (at university level) not representing the population was thus mitigated. The four classes were randomly divided into two groups (two classes formed a group) and assigned to either the intervention group (self-assessment) or the comparison group (peer assessment).

4.4.2.2 Lecturer participants

The two lecturer participants were experienced female English language lecturers of similar age (*Age* = 38 and 39 respectively) who taught Year Two in the selected university. Each of them had a Master's degree in English language and literature. They reported that they had no previous experience in using either self-assessment or peer-assessment in their writing classes. However, with more than ten years of tertiary teaching experiences, they were confident to ensure the smooth conduct of the intervention. The two lecturer participants were randomly allocated to either the intervention group or the comparison group classes.

4.4.3 Intervention fidelity and procedures

In this study, I measured the implementation fidelity from three perspectives. Firstly, to ensure the comparability between the intervention and comparison groups, both groups utilised the same textbook and writing tasks and received the same class instruction time and teacher feedback during the intervention.

Secondly, the quality of teaching delivery was ensured by teacher training prior the intervention. To be specific, I invited the EFL writing lecturer from the intervention group to attend a four-day training workshop with me (two hours per day) to upgrade her knowledge base of using self-assessment in EFL writing. Informed by previous studies (Andrade & Brown, 2016; Nielsen, 2019; Rolheiser, 1998), lesson plans for students' self-assessment development at different time points were developed (see Appendix F for a sample lesson plan). Both summative and formative uses of self-assessment were included in the intervention. The lecturer then implemented the four-month writing intervention to the intervention group exploring the effects of using self-

assessment on students' self-efficacy beliefs for self-assessment and writing performance over time. I made sure that the intervention group and comparison group teacher did not share their teaching practices, lesson plans, and other related materials during the time when the study was implemented and ongoing by asking them to sign an agreement (PIS form) in case they were influenced by each other. Besides that, after each monthly classroom observation, I discussed with the intervention group teacher briefly about the possible improvements in teaching practices, for example, if more students' self-assessment of writing practices or activities could be added in the next class so that the intervention quality was somehow ensured. Also, my observation of the both groups every month ensured the intervention group applied self-assessment while the comparison group did not. For the sake of equity and that students would not be disadvantaged through not taking the other assessment forms (either self- or peer-assessment), the same writing class (either using self- or peer-assessment), and resources used were provided for these students after the completion of the intervention programme.

Thirdly, the intervention fidelity was ensured through the verbal validation from students in their interviews, which were conducted at different time points as for my classroom observation. I checked with students regarding the lecturer's teaching practices in classes so that I could also learn what really happened in lessons that I did not observe.

The teaching intervention was conducted once a week (1.5 hours) from 09/2017 to 12/2018 for over 16 weeks. A week before the new semester started, I piloted the questionnaire, the writing test (the same topic as would be used in the pre-test), and the self-assessment rubric with ten English-major sophomore students (non-participants)

from the research site to resolve possible difficulties in understanding the instruments. The instruments were further revised based on students' feedback to ensure they had an accurate understanding of the contents.

At the beginning of the first semester in Year Two, all the participants for Part One (the intervention group and the comparison group) were first invited to complete the *Self-efficacy beliefs about self-assessment of EFL writing* (SESAEFLW) questionnaire in class to collect their demographic information as well as their reported self-efficacy for self-assessment of writing. Students spent, on average, approximately 20 minutes completing the survey, after which I collected the questionnaires.

Next, students were invited to take a timed pre-writing test, writing a cause and effect essay on a given topic (45mins). When students finished the writing task, they were asked to self-assess their writing. During the students' self-assessment task, both summative self-assessment and formative self-assessment were applied. For the summative self-assessment, students used the self-assessment writing rubric provided grading their writing texts from Level One to Level Five; for formative self-assessment, students identified, and commented on, the strengths and weaknesses in their writing text. I then collected students' essays and self-assessment statements. The data from these tests constituted the pre-test, serving as a baseline for further comparison.

To obtain an accurate estimation of students' writing performance, rater-training prior to scoring was essential. In this research, two independent raters, who are Ph.D. students majoring in Applied Linguistics at the University of Auckland with previous English teaching experiences in China, were invited to a training session before assessing students' writing texts in the pre-and post-tests. To avoid potential bias in essay evaluation, the raters were blind to the research conditions and were unfamiliar with the

research design. The training session covered the scoring rubric, as well as the rating process. Inter-rater agreement (reliability) was firstly checked in the trial scoring, with a random sample of 40 students' writing texts (forming approximately 20% of the entire data set) distributed to two raters. The two raters compared their scores and discussed any discrepancies in their evaluation in respect to the scoring criteria. The two raters started to score the rest of the writing samples independently only when their inter-rater reliability coefficients reached a satisfactory level ($r = .89$, $p < .001$), which ensured the reliability of the rating results (Smagorinsky, 2008).

During the intervention, students in both groups were required to perform self or peer-assessment on their assigned in-and-after class writing tasks every week. Students had opportunities to use the feedback from self or peer-assessment to improve their writing before the work was formally evaluated. Students' self or peer-assessment was not included in their final grade. Teachers agreed not to look at students' in-class self or peer-assessment, because keeping the act of self or peer-assessment partially private to students may promote willing and honest assessment practices (Andrade & Brown, 2016; Harris & Brown, 2018).

After the 16-week intervention, students from both groups were invited to complete the post-test questionnaire, the writing task and to self-assess their own writing texts as they did in the pre-test. Table 4.3 displays an overview of the intervention.

Table 4. 3 An Overview of the Intervention and the Comparison Group

Group	Participants	Instructor	N	Course Frequency	Content
Intervention Group (Self-assessment)	Undergraduate English major (Year Two)	One trained EFL lecturer	51	Once a week (1.5 hours for 16 weeks)	Regular writing course integrated with self-assessment-based writing instruction
Comparison Group (Peer-assessment)	Undergraduate English major (Year Two)	One EFL lecturer	41	Once a week (1.5 hours for 16 weeks)	Regular writing course integrated with peer-assessment based writing instruction

4.4.4 Data analysis

Before subjecting the data to inferential statistical analyses (e.g., independent and paired samples *t*-tests, and Pearson-correlation), data generated from the pre-and post-questionnaires, writing tests, students' self-assessment tasks were screened and cleaned to ensure response validity. Writing samples from the writing tests were transcribed into word-processor files before scoring. By doing so, the essay format consistency was ensured and possible rater bias with regard to handwriting was minimised (Powers, Fowles, Farnum, & Ramsey, 1994). I then further analysed the graded writing texts in terms of complexity, accuracy, and fluency (see the following sections for detailed description of the measures used).

The writing test data, together with students' self-assessment of writing outcomes, were entered into a database in SPSS 25, a statistical analysis software programme for Windows. Descriptive analysis (e.g., normality, mean score and standard deviation) and inferential analyses were checked to identify any discrepancies in the students' questionnaire, writing performance, students' and raters' assessment results before and after the intervention.

4.4.4.1 Independent samples t-tests and paired samples t-tests

A number of independent samples *t*-tests were conducted to compare whether there were significant differences between the means of self-assessment and peer-assessment on students' self-efficacy for self-assessment of writing, writing performance, linguistic development and rating accuracy in the pre- and the post-tests. In this stage, the independent variable was each individual condition (self-assessment or peer-assessment) and dependent variables were:

- Self-efficacy (writing tasks self-assessment self-efficacy and writing skills self-assessment self-efficacy).

- Writing performance (task achievement, coherence and cohesion, language resources, mechanics and overall performance).

- Linguistic development (complexity, accuracy and fluency).

- Students' rating accuracy.

The normal distribution of dependent variables was examined; no missing values or outliers were detected. A series of paired samples *t*-tests were conducted to investigate the possible changes in self-efficacy, writing performance, linguistic development, rating accuracy within each group from the pre- to the post-test. To control the effect that multiple *t*-tests have on the error rate, .05 was not used as the critical level of significance. Instead, the Bonferroni value was adjusted by the number of *t*-tests conducted (Field, 2013). In addition, to multiply .05 by the number of different comparisons (on the same data set), the benchmark value of Cohen's *d* was adopted to interpret the effect size (small = .2; medium = .5; large = .8) of the independent variable on dependent variables (Cohen, 1988, 1992).

4.4.4.2 Pearson product-moment correlations

The magnitude of the two variables' linear association was measured using the Pearson correlation coefficient (*r*) (Field, 2013). In Part One, Pearson correlation was applied to investigate (a) how similar the scores of students' and the raters' were in students' written texts in the pre- and post-tests (b) how students' self-reported efficacy for self-assessment of writing is associated with their actual writing performance in the pre- and post-tests. The effect sizes of the Pearson correlation coefficient (*r*) are small = .1, medium = .3, large = .5 (Cohen, 1988). As explained earlier, two raters were recruited for essay scoring, and the Pearson product-moment correlation coefficient test was used to calculate the inter-rater reliability of the two raters for this task. Pearson Product Moment correlation coefficient was also applied to check intra-rater reliability for data coding. High correlations represent high reliability.

4.4.4.3 Writing production measure—Complexity

Linguistic complexity concerns "the elaboration or ambition of the language which is produced" (Skehan, 1996, p. 23) and the necessity to integrate syntactic complexity and lexical complexity has been pointed out in previous studies (Lu, 2017; Skehan, 2009). This study applied a Web-based Lexical and Syntactic Complexity Analyzer (Lu, 2011, 2012), an automated interface to analyse the complexity measures. I entered students' essays in plain text to the interface, and the indices of syntactic and lexical complexity were generated automatically.

Syntactical complexity

Syntactic complexity denotes the "range of forms that surface in language production and the degree of sophistication of such forms" (Ortega, 2003, p. 492) and reveals

learners' written production variations (Casal & Lee, 2019). Informed by Wolfe-Quintero, Inagaki, and Kim (1998), the current study adopted two measures, namely, *mean length of T-units* (MLT) and the *number of clauses per T-unit* (C/T) to measure syntactic complexity in learners' essays. The total number of words produced in an essay divided by the number of *T-units* calculates the *mean length of T*-unit; the total number of clauses divided by with the total number of T-units, provides *the number of clauses per T-unit* (C/T). The size of these two values indicates the syntactic complexity of a student's essay.

Lexical complexity

Lexical variety or richness, lexical density, and lexical sophistication are components of lexical complexity (Read, 2010). In this study, *lexical variety* and *lexical density* were computed. *Lexical variety* indicates the range and sophistication of vocabulary production reflected in students' written work (Wolfe-Quintero et al., 1998). Mean segmental type-token ratio (MSTTR) is a measure to reveal the average ratio of the number of different words (type) used and the total number of words produced (token) in segments (Ellis & Barkhuizen, 2005). The larger the ratio is, the less repetitive the students' texts. Lexical density refers to the proportion of lexical or content words (as compared to grammatical or function) words in the total number of words (Read, 2010; Lu, 2017). In this study, the higher the ratio, the more informative is the student's text.

4.4.4.4 Writing production measure—Accuracy

Accuracy means "how well the target language is produced in relation to the rule system of the target language" (Skehan, 1996b, p. 23) and, similarly, accuracy indicates learners' "ability to be free from errors while using language'' (Wolfe-Quintero et al., 1998, p. 33). This study investigated the effect of using self-assessment on the overall quality of

students' written production using a general measure of accuracy as "a generalized measure of accuracy is more sensitive to detect differences between experimental conditions" (Skehan & Foster, 1997, p. 229). The general accuracy measures (as compared to specific linguistic structures) used for the current study, calculated the percentage of error-free clauses and the number of errors per 100 words.

The calculation of the percentage of error-free clauses is realised by the number of error-free clauses divided by the total number of clauses, multiplied by 100. Syntactical and morphological errors were marked. Errors in word order and words which were "nonexistent in English, or indisputably inappropriate" (Skehan & Foster, 1997, p. 195) were also counted, but errors in spelling and punctuation were ignored (Polio, 1997).

The number of errors per 100 words is achieved by dividing the total number of errors by the total number of words and then multiplying by 100. Such a measure takes account of every error in the written text and therefore resolves the problem that some clauses may include more than one error when measured by the percentage of error-free clauses (Foster & Skehan, 1996; Romano, 2019; Skehan, 2009). I coded the errors twice at different time points to ensure reliability.

4.4.4.5 Writing production measure—Fluency

To measure students' writing fluency, this study used the number of words produced per minute, calculated by the total number of words divided by the total amount of time devoted to writing (Abdel Latif, 2013; Kellogg, 1990). To establish the word count, I relied first on the suggestion in the Web-based Lexical and Syntactic Complexity Analyzer (Lu, 2011, 2012), and then I modified the word count following Polio's (1997)

instructions, in which, for example, a contraction and numbers were considered as one word, and title words were not included.

4.5 Part Two –Multiple-Case Study

A multiple case study connects a number of cases together to maximise the understanding of a complex phenomenon and can yield abundant data (Dörnyei, 2007; Merriam & Tisdell, 2016). In Part Two, to solicit in-depth information on the self-assessment intervention effects, I invited six students, with varied English proficiencies from the intervention group to form six case studies, which included attending semi-structured interviews and completing learning journals. Data collected via interviews and journals helped to gauge how the self-assessment of writing intervention affected their self-efficacy for self-assessment of writing, writing performance, writing linguistic development, rating accuracy of writing, and perception of self-assessment of writing. The classroom observations I made of the six students also helped to interpret their thoughts indicated in either interviews or journals. By addressing the following research questions, the qualitative data also served to verify and complement quantitative results from Part One.

a) What effects does the use of self-assessment have on students' (e.g., basic/proficient/advanced English proficiency) self-efficacy for self-assessment of EFL writing, student rating accuracy, and writing performance progress as well as their linguistic improvements in terms of complexity, accuracy, and fluency?

b) What are students' (e.g., basic/intermediate/advanced English proficiency) perceptions of, and attitudes to, self-assessment of writing, and the rubric used for self-assessment of writing?

c) What factors influence EFL students' self-assessment of writing practices?

4.5.1 Instrument

4.5.1.1 Semi-structured student interviews

Semi-structured interviews were conducted, initially, to reveal how students' self-efficacy for self-assessment of writing fluctuated and affected their engagement in the intervention. Secondly, the interviews allowed me to develop a broader sense of the factors that influenced self-assessment of writing practices, and to confirm, based on previous information achieved via other instruments, uncertainties in responses from the six students. In this study, interview schedules were developed and then employed at two points, prior to the intervention and one week after the intervention.

Participants were invited to attend the interviews individually for around 40 minutes in an empty classroom during the students' non-class time. Following the suggestions of Creswell and Creswell (2018) about designing interview protocols, the interview questions concerned students' self-efficacy for self-assessment of writing, writing performance and linguistic development, and perceptions and experiences in self-assessment practices. To avoid ambiguity and to facilitate students' expression of opinions, interviews were carried out in the students' first language, Mandarin, and audio-recorded with their consent. I took handwritten notes of significant non-verbal indicators. Appendix G shows interview questions asked prior to and after the intervention.

4.5.1.2 Learning journals

Participants were asked to keep structured learning journals in English from the start to the end of the intervention, that is, in the 4th, 10th and 16th week of the semester. Encouraged by the prompts, students recorded their attitudes to and understanding about

self-assessment of writing process, as well as possible changes, challenges and achievements they had experienced in the previous period. Each participant completed three writing journals within the provided learning journal frame with pen and paper in their own time (see Appendix H for details of the learning journal frame).

4.5.1.3 Raters' and students' comments on writing strengths and weaknesses

Through the pre- and post-writing tests, both raters and students indicated one writing dimension as the strength, and one writing dimension as the weakness in the writing; comments on why they chose a certain dimension as the strength or weakness were the last part of the self-assessment writing rubric. The six participants' pre- and post-comments were compared, respectively, with the raters' comments for possible overlaps or divergence.

4.5.2 Participants

Six students (two males and four females) from the intervention group were invited to serve as case studies. Purposive sampling (Creswell & Creswell, 2018) was used to obtain a range of levels of ability as it was hypothesised that experience and perception of self-assessment may vary with ability. In this study, purposive sampling also helps "enable detailed exploration and understanding of the central themes and puzzles which the researcher wishes to study" (Ritchie, Lewis, Nicholls, & Ormston, 2014, p. 73). Informed by Yoon and Lee's (2013) study investigating the effect of self-assessment on different proficiency levels, the participants (*M age* = 19, *SD* = .89) came from varied proficiency levels (basic/intermediate/advanced). Specifically, students were first divided into three English proficiency levels (basic/intermediate/advanced) based on their performance in the College entrance English exam, the English final exams in the

last semester, as well as the pre-test. For example, students with scores ranging from 60 to 69, 70 to 80, 80-100 were considered as basic learners, intermediate learners, and advanced learners respectively. Two students were randomly selected from each proficiency level to ensure a range of abilities was represented. Students' reported 10.2 years of learning English on average; none of the students had overseas learning experiences.

Table 4. 4 Participants Background Information

Name	Gender	Age	Previous Experience of Self-assessment of Writing	English Proficiency
Mia	Female	18	None	Advanced
Yvonne	Female	19	None	Advanced
Luke	Male	20	None	Intermediate
Joe	Male	19	None	Intermediate
Zoey	Female	18	None	Basic
Grace	Female	20	None	Basic

Note. Pseudonyms are used for students.

4.5.3 Procedures

Before and after the intervention, semi-structured interviews were conducted with the six students to explore any differences in their perceptions of their self-efficacy for self-assessment of writing, their writing performance and, particularly, their adaptation to, and experience of the intervention process. Interviews, conducted in Mandarin the interviewees' first language, lasted approximately 30 minutes; I audio-recorded, and transcribed verbatim, the interview responses. Students were first invited to check the interview transcription in Mandarin, and later, I showed them my English translation of

their transcripts. Students were aware that they could make any changes without explaining the reason. The same procedures were followed in the post interview at the end of the intervention, specifically, in the week after the post-test, again to ensure the trustworthiness of the data (Dörnyei (2007; Merriam & Tisdell, 2016). I collected students' comments on their writing strengths and weaknesses from the pre- and post-writing tasks, and also collected students' copies of journals each time when they completed them. No participants amended or withdrew any transcripts, translation or learning journals after four weeks of the intervention completion. All the collected data, therefore, were subjected to further analysis.

4.5.4 Data analysis

Trustworthiness refers to "a set of standards that demonstrates that a research study has been conducted competently and ethically" (Rallis & Rossman, 2009, p. 264). To establish the trustworthiness of inference for the qualitative aspect of this study, I followed the criteria introduced by Lincoln and Guba (1985), and further elaborated by Nowell, Norris, White, and Moules (2017), to conduct a trustworthy thematic analysis of the qualitative data. The criteria for trustworthiness included credibility, transferability, dependability and confirmability (Nowell et al., 2017). Initially, to ensure competent practices in this study, several techniques suggested by Lincoln and Guba (1985) were employed to establish its credibility. For instance, the qualitative data were gathered over a semester at different time points to reveal a relatively complete journey of the students' self-assessment of writing. Furthermore, different data collection methods, such as semi-structured interviews, learning journals, and classroom observations were used to investigate the research topic, and that allowed for further data triangulation and corroboration.

I also shared my transcripts, translation, codes, and analyses with the student participants for member checking to ensure that possible discrepancies in our interpretations were dealt with appropriately. With regard to transferability, I have endeavoured in Chapter Six to provide thick and faithful descriptions of students' responses to increase the possibility of transferring the findings to another similar context (Creswell & Creswell, 2018; Lincoln & Guba, 1985). To realise the dependability and confirmability of the study, I ensured that the qualitative research process was conducted in a clear and logical manner, with detailed explanations for the choices of research methods, instruments, participants, and data analysis approaches (Koch, 2006; Rallis & Rossman, 2009).

The six participants' answers to the interview and journal questions were analysed following a deductive, "top-down" thematic analysis approach, which is not only flexible but also advantageous in identifying, analysing and reporting themes from rich qualitative data (Braun & Clarke, 2006; Braun, Clarke, Hayfield, & Terry, 2019; King, 2004; Nowell et al., 2017). In the context of this study, themes stand for data categories related to the research questions (e.g., self-efficacy for self-assessment of writing or writing performance); and subthemes represent specific aspects (e.g., reasons of self-efficacy advancement, or writing accuracy improvement) that belonged to that category. Students' quotes from interviews and learning journals were selected according to the representativeness of their category.

Braun and Clarke (2006) have claimed that a "top-down" deductive approach affords "less a rich description of the data overall but more detailed analysis of some aspect of the data" (p. 84); data to answer the research questions were identified using such an approach. To be specific, with the help of the relevant theories and frameworks that underpin this study, I followed six phases that describe how the emerging themes were

identified, improved, and settled. Table 4.5 from Braun and Clarke (2006, p. 35), shows the six phases in detail.

In the first instance, I transcribed data generated from the 12 semi-structured interview sessions (audio-recordings); I then sent the transcripts to the corresponding participants for verification and amendment (participants could add to or cross out the statements which they made previously if their intended meaning was not delivered). To report the findings, after the transcripts were revised and finalised, I translated them from Mandarin to English. To verify the authenticity of data translation and interpretation, member checking was then conducted. Specifically, I invited two bilinguals with high Mandarin and English proficiency; one translated the transcripts from English to Mandarin, and the other one translated those transcripts back from Mandarin to English.

Table 4. 5 Six Phases of Thematic Analysis

Phase	Description of the process
1. Familiarising yourself with your data	Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas.
2. Generating initial codes:	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. Searching for themes:	Collating codes into potential themes, gathering all data relevant to each potential theme.
4. Reviewing themes	Checking in the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.
5. Defining and naming themes:	Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells; generating clear definitions and names for each theme.
6. Producing the report	The final opportunity for analysis. Selection of vivid, compelling extract examples, the final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.

Note. From Braun and Clarke (2006, p. 35)

Next, I went through the data recursively to achieve a holistic understanding of possible changes in students' self-efficacy for self-assessment, writing performance, linguistic development and their adaptation to the self-assessment intervention. I created a Microsoft Excel spreadsheet to input all the collated data of similar features (initial codes) and grouped them under the main themes and subthemes investigated in the research questions. I also used NVivo 12.0, a qualitative data processing software, to check and cross-check the word occurrence frequency, which implies the emergence of possible subthemes (from semi-structured interviews and learning journals) during the self-assessment intervention. After themes were generated, I reviewed them to ensure that there was no overlap or redundancy and the themes "tell a coherent, insightful story" about the data regarding the research questions (Braun et al., 2019, p. 12). To ensure the trustworthiness of coding, I coded the data twice, with two months intervening; also, a subset, which is approximately 30% of the total transcripts was randomly selected and coded by the two recruited raters (see details in section 4.4.3) to check for intra-coder and inter-coder reliability.

Raters' and students' comments on students' writing strengths and weaknesses were firstly compared for similarities and differences. I, then, referred to Hattie and Timperley's (2007) four levels of feedback to analyse how students' comments in explaining their writing strengths and weaknesses may have developed.

4.6 Chapter Summary

Starting with an overview of the research design, research paradigm, participants and instruments information, this chapter articulated the ethical considerations for the study. The third section detailed the instrument development and validation process in the preparatory stage. The succeeding two sections explained the two parts (Part One and

Part Two) of the main study by thoroughly focusing on instruments employed, participants involved, data collection procedures and data analysis. Part One examined the effect of using self-assessment on a range of variables, including students' self-efficacy for self-assessment of writing, writing performance and linguistic development, as well as students' rating accuracy. Part Two not only investigated how the self-assessment writing intervention impacted students on the above-mentioned variables, but also revealed students' learning and attitudinal adaptation to the intervention process and factors influencing students' self-assessment of writing practices. Data were gathered via questionnaires, writing tests, students' self-assessment tasks (ratings), semi-structured interviews, learning journals, and classroom observations. Different types of data (quantitative and qualitative) complemented and triangulated each other to address the research questions. Quantitative results and qualitative findings gathered from various instruments are summarised in the following two chapters.

CHAPTER FIVE

QUANTITATIVE RESULTS

This chapter first describes the instrument validation procedure and it then presents the comparison results between the intervention group and the comparison group in the pre-test. This is followed by the effects of the self-assessment intervention, on students' self-efficacy for self-assessment of writing and rating accuracy in comparison with the comparison group. The next section reports the effects of the intervention on students' writing performance and writing development, as evidenced in linguistic features such as complexity, accuracy and fluency. The fifth section describes the correlations between self-efficacy for self-assessment of writing and writing performance. The chapter concludes with a brief discussion of the implications of the quantitative data for the qualitative study, which will be presented in Chapter Six.

5.1 Instrument Validation: The Self-efficacy for Self-assessment of EFL Writing Questionnaire

5.1.1 Descriptive statistics

In cleaning the data, six respondents who gave the same rating for every item in the questionnaire were removed; no missing values were detected. As outliers, extreme values that fall a considerable distance from other values, need to be removed to satisfy normality (Tabachnick & Fidell, 2012). Eleven outliers were eliminated from the dataset. Descriptive statistical analyses indicated that the mean scores of the 25 items ranged from 2.97 (Item 16) to 4.19 (Item 4), with standard deviations ranging from .82 to .96. The skewness and kurtosis values, between -.97 to .23 and -.39 to .34, respectively (see

Table 5.1 for details), were within the ranges of cut-off values of +/-3.0 and +/-8.0

suggested by Kline (2016) for skewness and kurtosis. Assumptions of normality,

linearity, and homogeneity were not violated.

Table 5. 1 Descriptive Statistics of the Self-efficacy for Self-assessment of EFL Writing Questionnaire (SESAEFLWQ) (25 items)

Items	M	SD	Skewness	Kurtosis
Item 1-I can identify incorrect words' spellings in my writing or other writing samples.	3.35	.94	-.17	-.24
Item 2-I can assess if I use nouns, verbs, adjectives, etc. in the right way in my writing.	3.59	.94	-.42	-.01
Item 3-I can assess if I use plurals, verb tenses, prefixes, and suffixes in the right way in my writing.	3.34	.94	-.14	-.34
Item 4- I can assess whether I write a simple sentence with the right grammar or not.	4.18	.93	-.97	.36
Item 5- I can assess whether I write compound and complex sentences with the right grammar or not.	3.22	.93	-.19	-.18
Item 6- I can assess whether I write a logical paragraph with the topic sentence and supporting ideas or not.	3.13	.95	.03	-.17
Item 7-I can evaluate if my ideas are suitable to support my writing.	3.24	.84	-.02	-.14
Item 8-I can assess whether there is a clear organisation (e.g., ideas in order, effective linking words, etc.) in my writing.	3.13	.87	-.12	.20
Item 9-I can point out my strengths and weaknesses in my writing.	3.30	.95	-.17	-.39
Item 10- I assess whether I have clear goals before writing or not.	3.34	.90	-.08	-.33
Item 11- I can assess whether I realise my goals to improve my writing using different methods or not.	2.99	.85	.23	-.02
Item 12-I can revise my writing with the feedback from myself.	3.04	.96	-.06	-.34
Item 13-I can revise my writing when the teacher tells me what to do with my self-assessment.	3.68	.87	-.31	-.20
Item 14- I can self-assess my writing nicely with a detailed writing standard.	3.20	.86	-.06	-.22
Item 15- I can self-assess my writing better with a writing standard when writing samples in diverse levels are provided.	3.20	.87	-.13	-.28
Item 16-. I can self-assess my writing according to the descriptions in the writing standard of different writing styles.	2.97	.89	.07	-.29
Item 17-I can evaluate whether I achieve my goals in writing.	3.24	.87	-.15	-.19
Item 18-I can self-assess my writing process and point out the steps that I am not satisfied with.	3.55	.87	-.31	-.27
Item 19-I can estimate my writing performance correctly in class practices, homework, and exams.	3.00	.84	.02	-.19
Item 20- I can self-assess my writing with different methods if I were taught about them in the writing course.	3.25	.87	.13	-.34
Item 21-I can self-assess my writing well if I learn how to do writing self-assessment in class.	3.39	.85	-.12	-.21
Item 22-I can self-assess my writing well if I practice writing self-assessment regularly in class.	3.59	.82	-.25	-.05
Item 23- I can self-assess my writing better if I develop the writing standard on my own.	3.36	.83	-.14	-.04
Item 24- I can master writing self-assessment skills and self-assess my writing efficiently if my teacher encourages and supports me all the time.	3.61	.90	-.23	-.29
Item 25- I can improve my learning in other aspects of the English language by practicing self-assessment in English writing.	3.54	.85	-.26	.01

A sample size of 545 participants responded to the 25-item scale. As the structure of

self-efficacy for self-assessment of EFL writing is unclear in the existing literature, two

data sets with 245 and 300 participants were subjected to exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), respectively. As the above subjects-to-variables ratio exceeded the preferred ratio (5:1) suggested by Field (2013), the proposed questionnaire was suitable for factor analysis.

5.1.2 Exploratory factor analysis (EFA) results

The Kaiser-Meyer-Olkin measure indicated that the sample size was adequate with KMO = .92. Results of Bartlett's test of sphericity (df = 190, p < .001) were significant, meaning that the correlation among 25 questionnaire items was sufficiently large for an EFA. A maximum likelihood (ML) estimation by means of oblique rotation with Kaiser Normalization was performed on the 25 items as it "allows for the computation of a wide range of indexes of the goodness of fit of the model, permits statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals" (Fabrigar et al., 1999, p. 277). The parallel analysis was used because of its accurate performance in maintaining the number of components (O'Connor, 2000). A total of five items (6, 7, 8, 15, and 25) were excluded from further analysis due to their low loadings and complex loadings, and 20 items were retained. The removed items were not able to construct a reliable factor with other items possibly because students held limited and varied understanding of the organization of writing or ideas (items 6-8) prior to the intervention and they misunderstood the wording. The loading of the 20 items was all higher than the recommendation criterion (> .30) (See Appendix I for the descriptive statistics of the final version of Self-efficacy for Self-assessment of EFL Writing Questionnaire-SESAEFLW, with 20 items). Finally, a three-factor model, with more than three indicators each was extracted, explaining 52% of the total variance. Table 5.2 shows the EFA results and internal consistencies or reliabilities.

Table 5. 2 Factor Loadings for Exploratory Factor Analysis and Internal Reliability of the Three-factor SESAEFLW (N=245)

Factor	Item	Factor Loading			α
		1	2	3	
Writing Tasks Self-assessment Self-efficacy	Item 9- I can point out my strengths and weaknesses in my writing.	.70			.89
	Item 10- I can describe my goals before writing.	.70			
	Item 11- I can realise my goals to improve my writing using different methods.	.65			
	Item 12- I can revise my writing with the feedback from myself.	.67			
	Item 13-. I can revise my writing when the teacher tells me what to do.	.40			
	Item 14- I can understand the descriptions in writing standard of different writing styles.	.42			
	Item 16- I can self-assess my writing nicely with a detailed writing standard.	.53			
	Item 17- I can evaluate whether I achieve my goals in writing.	.60			
	Item 18- I can self-assess my writing process and point out the steps that I am not satisfied with.	.62			
	Item 19- I can estimate my writing performance correctly in class practices, homework, and exams.	.49			
Writing Skills Self-assessment Self-efficacy	Item 1- I can identify incorrect words' spellings in my writing or other writing samples.		.68		.86
	Item 2- I can assess if I use nouns, verbs, adjectives, etc. in the right way in my writing.		.74		
	Item 3- I can identify if I use plurals, verb tenses, prefixes, and suffixes in the right way in my writing.		.85		
	Item 4- I can write a simple sentence with the right grammar.		.61		
	Item 5- I can write compound and complex sentences with the right grammar.		.69		
Learning Needs Self-assessment Self-efficacy	Item 20- I can master the writing self-assessment skills and strategies if I were taught about them in the writing course.			.46	.84
	Item 21- I can self-assess my writing well if I learn how to do writing self-assessment in class.			.63	
	Item 22- I can self-assess my writing well if I practice writing self-assessment regularly in class.			.85	
	Item 23- I can apply the writing standard to my writing if I develop the writing standard on my own.			.73	
	Item 24- I can master writing self-assessment skills and self-assess my writing efficiently if my lecturer encourages and supports me all the time.			.61	

Based on an examination of items gathering around each factor and the related literature, the three factors were named: Factor 1 Writing Tasks Self-assessment Self-efficacy (39% variance); Factor 2 Writing Skills Self-assessment Self-efficacy (8% variance); and Factor 3 Learning Needs Self-assessment Self-efficacy (5% variance). Cronbach's alpha

coefficient for each of the three factors was greater than the threshold value (.70), indicating satisfactory internal consistency reliability of the instrument (DeVellis, 2012).

5.1.3 Confirmatory factor analysis (CFA) results

There was no missing data in another sample set (N = 300) and concerning the sensitivity of structural equation modelling (SEM) to multivariate normality, Mahalanobis distance was examined to further remove 15 outliers using SPSS Statistics 25. To appraise the plausibility of the hypothesised three-factor structure generated in EFA results, the final sample size of 285 cases were submitted for a series of CFAs using IBM SPSS Amos 25. A zero-order correlation model was proposed, and the maximum likelihood (ML) method was used to estimate the model's parameters and fit indices. In this three-factor correlated model, each indicator only loaded on the factor it was designed to measure, and measurement errors linked with each indicator were uncorrelated. Factor covariances were free to be estimated. Non-zero loadings were allocated for each item-pair measure on the subcategories of self-efficacy and the questionnaire aimed to measure a zero loading on all other factors. Figure 5.1 shows the structure of the Three-Factor Model 1 and the calculation of zero-order correlations between the factors.

Various omnibus fit statistical analyses were applied to Model 1 for model fit evaluation. Generally, the primary fit of Model 1 exhibited an overall acceptable model fit (x^2 = 318.72; df = 167; $p < .001$; x^2/df = 1.91; RMSEA = .05; SRMR = .05; TLI = .91; CFI = .94). Although the above values met the desirable model fit criteria suggested by Hu and Bentler (1999), there was a high correlation (r = .85) between the factors of Writing tasks and learning needs self-assessment self-efficacy. This result indicated overlapping traits between the two factors, and that one of them should be dropped to ensure construct validity (Kline, 2016).

Figure 5. 1. The Three-factor Correlated Model of SESAEFLW

Note. Performance = Writing tasks self-assessment self-efficacy; Linguistic = Writing skills self-assessment self-efficacy; Needs = Learning needs self-assessment self-efficacy.

Therefore, Factor 3 Learning Needs Self-assessment Self-efficacy was removed, and its indicators were merged into Factor 1 for further analysis. CFAs were then performed to a two-factor model (Factor 1 Writing Tasks Self-assessment Self-efficacy; Factor 2 Writing Skills Self-assessment Self-efficacy) with the same sample group using the maximum likelihood (ML) method. Table 5.3 displays regression weights for this two-

factor correlated model of self-efficacy for self-assessment of EFL writing (SESAEFLW).
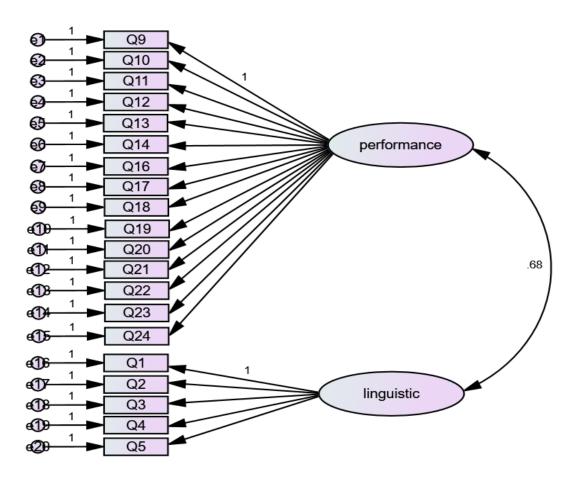
Table 5. 3 Regression Weights for the Two-factor Correlated Model of SESAEFLW

SESAEFLW	Unstandardised Estimate	Standardised Estimate	CR
Item 1-Linguistic	1.00[a]	.69	a
Item 2- Linguistic	1.11	.81	11.45[***]
Item 3- Linguistic	1.02	.73	10.59[***]
Item 4- Linguistic	.61	.49	7.47[***]
Item 5- Linguistic	.80	.63	9.30[***]
Item 9-Performance	1.00[a]	.55	a
Item 10-Peformance	1.01	.57	7.66[***]
Item 11-Performance	.98	.59	7.89[***]
Item 12-Performance	1.25	.64	8.29[***]
Item 13-Performance	1.17	.69	8.69[***]
Item 14-Performance	.99	.60	7.92[***]
Item 16-Performance	1.03	.62	8.12[***]
Item 17-Performance	1.01	.63	8.18[***]
Item 18-Peformance	1.05	.63	8.20[***]
Item 19-Performance	1.06	.61	8.07[***]
Item 20-Performance	1.14	.67	8.51[***]
Item 21-Performance	1.39	.81	9.48[***]
Item 22-Performance	1.26	.77	9.24[***]
Item 23-Performance	1.23	.74	9.04[***]
Item 24-Performance	1.27	.69	8.66[***]

Note: [***] = $p < .001$; "a" means the regression weight was fixed at 1.00 for model identification and no critical ratio was computed; Linguistic = Writing Skills Self-assessment Self-efficacy; Performance = Writing Tasks Self-assessment Self-efficacy

CFA results for the above two-factor model ($x^2 = 400.97$; $df = 169$; $p < .001$; $x^2/df = 2.37$; RMSEA = .07; SRMR = .04; TLI = .90; CFI = .91) suggested that the model fit indices were generally satisfactory. The 20-item parameter estimates were significant statistically at $p < .001$. The effect size was relatively large as standardised estimates loadings on the postulated latent constructs were mostly higher than the benchmark value of .50 (Raykov & Marcoulides, 2008).

Figure 5. 2. The Two-factor Correlated Model of SESAEFLW

Using Cronbach's Alpha, the results showed that the reliability values for the two factors were .92 and .80 for Factor One and Two respectively, supporting a satisfactory reliability (≥ .70) for internal consistency. The results also indicated that the two factors correlated moderately with one another, and at the same time, shared enough distinction (see Figure 5.2), which confirmed the discriminant validity. Compared with Model 1, the three-factor model, this two-factor model seemed to have a better fit and was therefore retained as the final model.

5.1.4 Summary

The hypothesised structure of EFA, further revised with the evidence provided by CFA validation procedures for this instrument, led to a two-factor model of self-efficacy for self-assessment of EFL writing, including writing tasks and writing skills self-assessment self-efficacy. The medium correlation between the two dimensions meant that they were distinct from one another but correlated under the same construct (Kline, 2016). Satisfactory results from factor analyses have provided evidence for the structural aspect of construct validity.

5.1.4.1 Writing tasks self-assessment self-efficacy

The first dimension was labelled *writing tasks self-assessment self-efficacy* (Shell et al., 1989), describing students' judgements regarding their capability to self-assess their performance as well as learning needs in EFL writing (e.g., I can point out my strengths and weaknesses in my writing). Bandura (2006) argues that writing self-efficacy evaluation should "be linked to the behavioural factors over which people can exercise some control" (p. 310). Hence, data collected from writing tasks self-assessment self-efficacy would offer insight into how students' beliefs in their capabilities influence their behaviour in the specific domain of self-assessment of EFL writing.

5.1.4.2 Writing skills self-assessment self-efficacy

The second dimension was named *writing skills self-assessment self-efficacy*, referring to students' judgements of their capability in correctly applying various lexical, syntactic, and mechanical skills in their writing (e.g., I can assess whether I write compound and complex sentences with the right grammar or not). Whereas many studies have acknowledged the potential effect of self-efficacy in writing skills on L1 writing

performance (Bruning et al., 2013), inadequate attention has been paid to EFL writing contexts. It is anticipated that findings afford a better understanding of the interaction of students' confidence in assessing their linguistic competence and EFL writing performance itself.

5.2 Comparisons of the Intervention and the Comparison Groups in the Pre-test

This part evaluated the potential differences concerning social factors, self-efficacy for self-assessment of writing, writing performance, as well as students' rating accuracy in self-assessment of writing between the two groups in the pre-test, using independent samples t-tests. Assumptions of t-tests were satisfactorily met. To protect against Type I errors, the Bonferroni procedure was applied to each t-test with the adjusted alpha (.05 divided by the number of t-tests conducted, which should be equal to the number of all dependent variables) (Raykov & Marcoulides, 2008).

5.2.1 Students' backgrounds

A series of independent samples t-tests were applied to assess the social contextual differences (i.e., age and years of English learning) between the two conditions (self-assessment and peer-assessment). The adjusted Bonferroni value was at .0025. According to the descriptive analyses, students' average ages were M_{SA} =19.59 (SD =0.67) and M_{PA} =19.54 (SD = 0.67) in the intervention and comparison groups. Students from those two conditions reported 11 years of English learning experiences on average M_{SA} =11.55 (SD =0.99) and M_{PA} =11.84 (SD = 0.77). Results showed no significant differences between groups in the two conditions regarding age and years of learning English (see Table 5.4). Students' gender between two conditions was compared and

checked by a Chi-square test of independence, with no significant difference detected,

x^2 (1) = 1.09, p = .30.

Table 5. 4 Descriptive Statistics and Results of Independent Samples t-tests of Age and
Years of English Learning between the Two Groups (Intervention and Comparison)

Variables	Condition	N	M	SD	t	p	95% CI	
							LL	UL
Age	SA	51	19.59	.67	.37	.72	-.23	.33
	PA	41	19.54	.67				
Year of English Learning	SA	51	11.55	.99	-1.49	.14	-.65	.09
	PA	41	11.84	.77				

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

5.2.2 Self-efficacy for self-assessment of writing

With adjusted Bonferroni value at .0025, a series of independent samples t-tests were
performed on the two groups (intervention and comparison) to examine if there were
prior differences between the two groups concerning self-efficacy for self-assessment of
writing. According to Table 5.5, the two groups showed no significant difference in their
self-efficacy for self-assessment of writing in either the writing tasks or the writing skills
aspect in the pre-test. This ensured that the two groups were compared on an equal
footing when the intervention effects were to be checked later on in the main study.

Table 5. 5 Descriptive Statistics and Independent Samples *t*-tests of Self-efficacy between the Two Conditions Self-assessment (SA) and Peer-assessment (PA) in the Pre-test

Self-efficacy	Condition	N	M	SD	t	p	95% CI LL	95% CI UL
Writing Tasks Self-assessment	SA	51	2.61	.74	.17	.47	-.29	.34
	PA	41	2.59	.78				
Writing Skills Self-assessment	SA	51	2.90	.84	-.02	.70	-.41	.34
	PA	41	2.90	.78				

Note. This is a 5-point Likert scale; 1 =not at all sure I can do; 2 = slightly sure I can do; 3 = moderately sure I can do; 4 = very sure I can do; 5 = completely sure I can do; *CI* = confidence interval; *LL* = lower limit; *UL* = upper limit.

5.2.3 Writing performance

As shown below in Table 5.6, the conduct of independent samples *t*-tests revealed no significant difference in the pre-test writing scores between the intervention group and the comparison group with adjusted Bonferroni value at .001. (self-assessment and peer-assessment) before the intervention (see Table 5.6).

Table 5. 6 Descriptive Statistics and Independent Samples *t*-tests of Writing Test Scores Between the Conditions of Self-assessment (SA) and Peer-assessment (PA) in the Pre-test

Writing Test score	Condition	N	M	SD	t	p	95% CI LL	95% CI UL
Task Achievement	SA	51	3.18	.51	1.19	.24	-.09	.37
	PA	41	3.04	.60				
Coherence and Cohesion	SA	51	3.17	.51	.38	.71	-.16	.23
	PA	41	3.13	.40				
Language Resources	SA	51	3.07	.48	-.49	.63	-.26	.16
	PA	41	3.12	.52				
Mechanics	SA	51	2.89	.52	.67	.51	-.13	.26
	PA	41	2.82	.40				
Total Scores	SA	51	12.30	1.55	.62	.54	-.43	.81
	PA	41	12.11	1.37				

5.2.4 Students' rating accuracy in self-assessment of writing

To measure students' initial consistency in self-assessment of writing, the correlation between the students' rating and the raters' rating on writing samples from the pre-test was applied. To be specific, how the students' rating correlated with the raters' rating in four individual dimensions (i.e., task achievement, coherence and cohesion, language resources, and mechanics) as well as the overall performance of EFL writing was calculated. Pearson correlation coefficient (*r*), as an indicator of the linear relationship between two variables (Field, 2013), was applied to reveal the strength of correlation for each dimension, respectively. Results of the students' and raters' correlation before the intervention (self-assessment) using bivariate Pearson correlations with two-tailed significance can be found in Table 5.7. The effect size of the Pearson correlation coefficient *r* was reported referring to Cohen (1988, 1992), in which $r = .1$, $r = .3$, and $r = .5$ representing small, medium and large effect sizes respectively.

Table 5. 7 Correlations of the Students' and the Raters' Rating on Pre-writing Test Samples by Two Groups

Variables	TA	CC	LR	M	Overall
	Effect Size *r*-square				
	Pre	Pre	Pre	Pre	Pre
SR vs TR (SA)	.19	.26	.06	.24	.34**
SR vs TR (PA)	.23	.25	.15	.11	.45**

Note. TA = Task Achievement; CC = Coherence and Cohesion; LR = Language Resources; M = Mechanics; SR = students' rating; TR = teacher's rating; ** = $p < .01$; * = $p < .05$.

Bivariate correlations revealed that students' and the raters' ratings correlated positively for three of the four writing dimensions (the dimension of language resources was the exception), and for the overall performance of the pre-writing test samples. Notably, it was only on language resources that students and the raters' ratings in both groups were negatively related in the pre-test. Furthermore, students' ratings of both groups showed a low correlation with the raters' rating for all the other dimensions in the pre-test (from $r = .11$ to $r = .26$). In regard to the overall performance, students' and the raters' rating in both groups had significant correlations in the pre-writing test, that is, moderate correlation coefficients were found in two groups: Students' and the raters' rating correlation for the intervention (self-assessment) group was $r = .34$ and the comparison (peer-assessment) group was $r = .45$. In other words, in the pre-test, students in both groups were accurate at only a medium level in self-assessing their overall writing performance.

5.2.5 Summary

The self-assessment and peer-assessment group were comparable, statistically, for their social factors (i.e., age, years of English learning and gender), self-efficacy for self-assessment of writing (performance and linguistics dimensions), writing performance, as well as their rating consistency in self-assessment of writing prior to the intervention. Therefore, further analyses were feasible based on their comparable conditions.

5.3 Effects of Self-assessment-based Intervention on Students' Self-efficacy, and Rating Accuracy in Self-assessment of Writing

The research question this section deals with was: "What effects does the use of self-assessment in EFL writing classes have on students' self-efficacy for self-assessment of

EFL writing (SESAEFLW), and students' rating accuracy in self-assessment of EFL writing, as compared with raters' judgement?" To answer this question, initially, a number of independent and paired samples t-tests were used to assess potential changes in self-efficacy within and between the two groups. To protect against Type I errors, a Bonferroni adjustment was adopted to set the alpha level at each t-test (.05 divided by the number of t-tests conducted), which should be equal to the number of all dependent variables (Field, 2013; Raykov & Marcoulides, 2008). Students' rating consistency in self-assessment of writing was also measured by the Pearson correlation coefficient (r).

5.3.1 Self-efficacy for self-assessment of EFL writing

Students in both groups, with approximate medium self-efficacy for self-assessment of writing before the intervention, demonstrated a considerable increase after the intervention, with an adjusted Bonferroni value at .0025. As shown in Table 5.8, students from the intervention group made relatively stronger gains than the comparison group on the SESAEFLW at the post-test with Cohen's d estimated at 3.74 and 2.0 for self-efficacy in writing tasks and writing skills self-assessment respectively, much higher than Cohen's d guidelines (1988, small = .2, medium = .5, large = .8).

Table 5. 8 Descriptive Statistics and Results of Paired Samples *t*-tests of SESAEFLW in the Pre- and Post-tests in Two Groups – the Intervention Group and the Comparison Group

Self-efficacy	Group	Pre-test (T1)			Post-test (T2)			T1 vs. T2		
		N	M	SD	N	M	SD	*t*	*p*	Cohen's *d*
Writing Tasks Self-assessment	intervention	51	2.61	.74	51	3.79	.41	-14.29	< .001	3.74
	comparison	41	2.59	.78	41	2.91	.36	-26.69	< .001	1.56
Writing Skills Self-assessment	intervention	51	2.90	.84	51	3.82	.51	-9.06	< .001	2.00
	comparison	41	2.90	.78	41	3.19	.52	-9.99	< .001	1.42

The following independent samples *t*-tests results showed that there were significant differences between the two groups regarding self-efficacy for self-assessment of writing. The results suggested the positive effect of self-assessment-based writing intervention on efficacy for self-assessment in writing as the intervention group scored significantly higher than the comparison group in self-efficacy with strong effect sizes of Cohen's *d* = 2.67 and 1.51 for each dimension, respectively (see Table 5.9).

Table 5. 9 Descriptive Statistics and Independent Samples *t*-tests of Self-efficacy for Self-assessment of Writing after the Intervention

Self-efficacy	Condition	N	M	SD	*t*	*p*	95% *CI*	
							LL	*UL*
Writing Tasks Self-assessment	SA	51	3.79	.41	15.22	< .001	.95	1.24
	PA	41	2.69	.23				
Writing Skills Self-assessment	SA	51	3.81	.51	7.65	< .001	.57	.97
	PA	41	3.04	.45				

5.3.2 Students' rating accuracy in writing

With comparable accuracy in self-assessment of writing in the pre-test, the correlation between the students' rating and the raters' rating on writing samples from the post-test concerning four individual dimensions (i.e., task achievement, coherence and cohesion, language resources, and mechanics) as well as the overall performance of EFL writing was investigated. The strength of students' and raters' correlations in different dimensions in the post-writing test are displayed in Table 5.10 using bivariate Pearson correlations with two-tailed significance.

Table 5. 10 Correlations of Students' and the Raters' Rating on Post-writing Test Samples by the Intervention Group and the Comparison Group

Variables	TA		CC		LR		M		Overall	
	Effect Size *r*-square									
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
SR vs RR (Intervention)	.19	.38**	.26	.13	.06	.25	.24	.11	.34**	.38**
SR vs RR (Comparison)	.23	.12	.25	.13	.15	.29	.11	.01	.45**	.42**

Note. TA = Task Achievement; CC = Coherence and Cohesion; LR = Language Resources; M = Mechanics; SR = students' rating; RR = raters' rating; ** = $p < .01$; * = $p < .05$.

For both groups, according to the above bivariate correlations, the students' rating correlated with the raters' rating for the four writing dimensions and with the overall performance, positively in the post-test. It is notable that for the dimension of language resources, while the students' and the raters' ratings in both groups were negatively associated in the pre-test, the students' and the raters' ratings showed a positive correlation in the post-test, with a non-significant small effect (SA: $r = .25$; PA: $r = .29$).

The correlation between the students' and the raters' ratings for other dimensions decreased, but differently, for both groups in the post-test. An exception was task achievement for the intervention group, in which the students' and the raters' ratings were significantly positively correlated ($r = .38$). For the overall performance, similar to the pre-writing test, students' and the raters' ratings in both groups had significant moderate positive correlations at the post-writing test. Although remaining at a medium level, the students' and the raters' rating correlation for the intervention group indicated a small growth (from $r = .34$ to 38); in contrast, the comparison group had a slight decrease (from $r = .45$ to .42).

5.3.3 Summary

In summary, from the pre-test to the post-test, both groups showed a significant within-group increase in self-efficacy for self-assessment of writing after one semester. Correlations of the students' and the raters' ratings for both groups were significant only in terms of overall writing performance, and task achievement for the intervention group at the post-test. No significant associations were found in coherence and cohesion, language resources, and mechanics. The results imply students' low rating accuracy in terms of self-assessment of individual writing dimensions, but relatively higher, intermediate, rating accuracy in self-assessing their overall writing performance.

5.4 Effects of Self-assessment-based Intervention on Writing Performance and Linguistic Development

This section responds to the research question: "What effects does the use of self-assessment in EFL writing classes have on students' writing performance and writing linguistic improvement in terms of complexity, accuracy, and fluency?" The data used

were gathered from participants' writing samples in the pre-and post-tests. To check whether students' writing improved because of the use of self-assessment, I first examined students' writing quality from four dimensions, and examined students' writing performance in terms of their complexity, accuracy, and fluency of EFL written texts.

The present study adopted seven dependent variables based on writing performance measurements described in the previous literature to assess learners' writing performance (Lu, 2011, 2012; Polio, 1997), (see Table 5.11 for a summary and Chapter 4 for more information). The assumption of normality for the following dependent variables was satisfied and Bonferroni adjustment was applied in each *t*-test (0.5 divided by the number of *t*-tests conducted, which equals the number of all dependent variables) to avoid Type I errors (Raykov & Marcoulides, 2008).

Table 5. 11 Summary of Dependent Variables

Syntactic complexity	1	MLT	Mean length of T-units
	2	C/T	Number of clauses per T-unit
Lexical complexity	3	MSTTR	Mean segmental type-token ratio (lexical variation)
	4	LD	Lexical density (percentage of content words)
Accuracy measures	5	Error-free clauses	Percentage of error-free clauses
	6	Error/100	Error per 100 words
Fluency	7	WPM	Words per minute

5.4.1 Writing performance

It can be seen from Table 5.12 below that students from both the intervention (self-assessment) and the comparison group (peer-assessment) experienced an increase of writing scores in task achievement, coherence and cohesion, language resources, mechanics and overall performance. Adjusting the Bonferroni value at .001, and according to the Cohen's *d* levels shown below, the effect of improvement was very strong at the post-test for both groups. The intervention group outperformed the comparison group in task achievement, coherence and cohesion, mechanics, and overall writing scores with effects of different magnitude; students from the intervention group achieved greater significant gains (Cohen's $d = 1.89$) in their overall writing scores than the comparison group (Cohen's $d = .78$).

Table 5. 12 Descriptive Statistics and Results of Paired Samples *t*-tests of Writing Scores in the Pre- and Post-tests in the Intervention (SA) and the Comparison Group (PA)

Writing Dimensions	Group	Pre-test (T1)			Post-test (T2)			T1 vs. T2		
		N	M	SD	N	M	SD	*t*	*p*	Cohen's *d*
Task Achievement	SA	51	3.18	.51	51	4.02	.45	-11.68	< .001	1.62
	PA	41	3.04	.60	41	3.74	.42	-8.33	< .001	1.32
Coherence and Cohesion	SA	51	3.17	.51	51	3.96	.46	-11.08	< .001	1.54
	PA	41	3.13	.40	41	3.63	.44	-9.39	< .001	1.49
Language Resources	SA	51	3.07	.48	51	3.59	.50	-7.65	< .001	1.06
	PA	41	3.12	.52	41	3.59	.50	-6.49	< .001	1.03
Mechanics	SA	51	2.89	.52	51	3.43	.51	-7.65	< .001	1.06
	PA	41	2.82	.40	41	3.10	.43	-4.78	< .001	0.76
Total Score	SA	51	12.30	1.55	51	14.99	1.57	-13.63	< .001	1.89
	PA	41	12.11	1.37	41	13.77	2.26	-4.94	< .001	0.78

Furthermore, it is apparent that the effects of using self-assessment or peer-assessment on writing depend on the specific language domain examined. For example, in Table 5.13, the intervention group appears to have a larger effect on task achievement than coherence and cohesion. Moreover, it can be seen that results from independent samples *t*-tests demonstrate that the writing performance for the intervention group and the comparison group were mostly clearly differentiated (*p* < .001) in the post-test.

Specifically, the intervention group performed significantly better in the dimensions of task achievement, coherence and cohesion, and mechanics (*p* < .001) and showed superior performance in terms of their overall writing scores (*p* < .001). However, for language resources, although both groups registered within-group writing performance improvement (see Table 5.12) in the post-test, they did not exhibit any between-group differences (see Table 5.13).

Table 5. 13 Descriptive Statistics and Independent Samples *t*-tests of Writing Test Scores between the Intervention group (SA) and the Comparison group (PA) in the Post-test

Writing Test score	Condition	N	M	SD	*t*	*p*	95% *CI*	
							LL	*UL*
Task Achievement	SA	51	4.04	.43	3.59	.001	.15	.50
	PA	41	3.72	.43				
Coherence and Cohesion	SA	51	3.97	.46	3.65	.000	.16	.53
	PA	41	3.62	.44				
Language Resources	SA	51	3.58	.50	-.24	.810	-.23	.18
	PA	41	3.60	.50				
Mechanics	SA	51	3.44	.51	3.41	.001	.14	.54
	PA	41	3.10	.42				
Total Scores	SA	51	15.03	1.57	3.18	.002	.47	2.05
	PA	41	13.76	2.23				

Note. The writing rubric measured four dimensions of writing performance, using a five-scale scoring scheme (1-5) for each dimension. CI = confidence interval; LL = lower limit; UL = upper limit.

5.4.2 Results of complexity measures

The complexity of students' written texts in the pre- and post-tests were analysed in two dimensions, syntactic and lexical complexity, and, within each dimension, two measurements were selected.

5.4.2.1 Syntactic complexity

Table 5.14 presents the descriptive statistics as well as the independent samples *t*-test results for the two syntactic complexity measures. As is indicated in the table, the two groups, learners' performance in the mean length of T-units (MLT) and the number of clauses per T-unit (C/T) did not vary significantly from the pre-test to the post-test. Neither of the two groups made significant improvement over time.

Table 5. 14 Descriptive Statistics and Independent Samples *t*-tests of Syntactic
Complexity Measures in the Pre-and Post-tests

	Syntactic Complexity	Condition	N	M	SD	*t*	*p*	95% *CI* LL	UL
Pre-test	MLT	SA	51	13.13	2.82	-1.74	.05	-2.31	.02
		PA	41	14.21	3.12				
	C/T	SA	51	1.63	.30	-1.73	.05	-.24	1.94
		PA	41	1.74	.32				
Post-test	MLT	SA	51	13.86	2.11	-.25	.40	-1.18	.92
		PA	41	13.99	2.94				
	C/T	SA	51	1.51	.22	-.43	.34	-.13	-.08
		PA	41	1.53	.30				

Note. MLT = Mean length of T-units; C/T = Number of clauses per T-unit

A series of paired samples *t*-test was employed to examine whether there was potential change in either of the two syntactic complexity measures within each group. Based on the results from paired samples *t*-tests, there was no significant variation for either the self-assessment group (Pre-*M* = 13.13, *SD* = 2.82; Post-*M* = 13.86, *SD* = 2.11; *p* = .06) or the peer-assessment group (Pre-*M* = 14.21, *SD* =3.12; Post-*M* = 13.99, *SD* = 2.94; *p* = .33) in MLT from the pre-test to the post-test.

However, a significantly negative change was found in the number of clauses per T-units within each group between the pre-test and the post-test. Specifically, the number of clauses per T-unit in students' writing experienced a noticeable decrease over time for the self-assessment group (Pre-*M* = 1.63, *SD* = .30; Post-*M* = 1.51, *SD* = .22; *p* = .02) and the peer-assessment group (Pre-*M* = 1.74, *SD* = .32; Post-*M* = 1.53, *SD* = .30; *p* = .00). Although the effect size of the differences for the self-assessment group was small (Cohen's *d* = .32), the effect size for the peer-assessment group reached the medium level (Cohen's *d* = .58).

5.4.2.2 Lexical complexity

Measurements of lexical complexity comprise the mean segmental type-token ratio (MSTTR), namely, lexical variation and lexical density (the percentage of content words). Table 5.15 below shows a between groups comparison in the pre- and post-tests, in which the *p*-value indicated that there was no difference between the groups in the pre-test or in the post-test in the ratio of mean segmental type-token and lexical density.

Table 5. 15 Descriptive Statistics and Independent Samples *t*-tests between the Intervention Group and Comparison Group of Lexical Complexity Measures in the Pre-and Post-tests

	Lexical Complexity	Condition	N	M	SD	*t*	*P*	95% *CI* LL	UL
Pre-test	MSTTR	SA	51	.77	.03	1.12	.13	-.01	.02
		PA	41	.76	.34				
	LD	SA	51	.52	.38	-1.70	.07	-.28	1.94
		PA	41	.53	.04				
Post-test	MSTTR	SA	51	.77	.03	.97	.17	-.02	.01
		PA	41	.78	.04				
	LD	SA	51	.54	.03	-1.10	.14	-.02	.01
		PA	41	.55	.03				

Note. MSTTR = Mean segmental type-token ratio; LD = Lexical density

Table 5.16 presents within-group differences of lexical complexity, which show that learners in both the self-assessment and peer-assessment group had a similar slight increase in lexical density in the post-test. In the pre-test, the average mean of lexical density for the self-assessment group was .52 (*SD* = .04), whereas, in the post-test, the mean value increased to .54 (*SD* = .03, *t* = -3.59, *p* < .001) indicating the self-assessment intervention had a significant difference, a medium effect, on lexical density (Cohen's *d* = .57) in students' writing. For the peer-assessment group, however, the slight growth in lexical density (Pre-*M* = .53, *SD* = .04; Post-*M* = .54, *SD* = .03) was not statistically significant (*p* = .06 *t* = -1.62).

Table 5. 16 Paired Samples *t*-tests within the Intervention Group and Comparison Group of Lexical Complexity Measures in the Pre-and Post-tests

Lexical Complexity	Group	Pre-test (T1)			Post-test (T2)			T1 vs. T2		
		N	M	SD	N	M	SD	*t*	*p*	Cohen's *d*
MSTTR	SA	51	.77	.03	51	.77	.03	.66	.51	
	PA	41	.76	.04	41	.77	.04	-1.57	.12	
LD	SA	51	.52	.04	51	.54	.03	-3.59	< .001	.57
	PA	41	.53	.04	41	.54	.03	-1.62	.06	

5.4.3 Results of accuracy measures

Measures of error-free clauses and errors per 100 words, were employed to analyse students' written texts in the pre- and post-tests. Table 5.17 below displays the descriptive statistics and independent samples *t*-test results, which show that there were no significant differences between groups for the proportion of error-free clauses and the number of errors per 100 words ($t = -.86$, $p = .20$ and $t = 1.49$, p = .07, respectively) in the pre-test. In the post-test, however, the self-assessment group differed significantly from the peer-assessment group on the two accuracy measures ($p < .001$). Specifically, the magnitude of difference indicated that students in the intervention group performed better than those in the comparison group with error free clauses (Cohen's $d = .92$) as well as errors per 100 words (Cohen's $d = -.82$).

Table 5. 17 Descriptive Statistics and Independent Samples *t*-tests of Accuracy Measures in the Pre-and Post-tests

	Accuracy Measures	Condition	N	M	SD	t	p	95% CI	
								LL	UL
Pre-test	Error Free Clauses	SA	51	50.23	13.16	-.86	.20	-.83	3.29
		PA	41	52.76	14.89				
	Error/100	SA	51	9.03	2.78	1.49	.07	-.28	1.94
		PA	41	8.20	2.50				
Post-test	Error Free Clauses	SA	51	64.71	13.70	4.37	< .001	6.83	18.25
		PA	41	52.17	13.69				
	Error/100	SA	51	5.27	2.20	-3.92	< .001	-.2.72	-.89
		PA	41	7.07	2.19				

Note. *CI* = confidence interval; *LL* = lower limit; *UL* = upper limit.

Paired samples *t*-tests were then conducted to compare gains in accuracy measures in the post-test within each group. Results (see Table 5.18) revealed that both the intervention and the comparison groups had improved markedly on one of the accuracy measures. While students from the intervention group had significant gains in the percentage of error-free clauses (Pre-*M* = 50.23, *SD* = 13.16; Post-*M* = 64.71, *SD* = 13.70, *t* = -6.64, *p* < .001) with a large effect size (Cohen's *d* = -.92), for students in the comparison group, there was little difference in their proportion of error-free clauses between the pre- and post-tests (Pre-*M* = 52.76, *SD* = 14.89; Post-*M* = 52.17, *SD* = 13.69; *p* = .38).

Despite varying in gains in error-free clauses, both groups revealed a noticeable improvement in terms of their number of errors per 100 words (*p* < .001). Students from the intervention group (Pre-*M* = 9.03, *SD* = 2.78; Post-*M* = 5.27, *SD* = 2.20) made fewer errors than the comparison group (Pre-*M* = 8.20, *SD* = 2.50; Post-*M* = 7.07, *SD* = 2.19)

in the post-test. The effect size of the gains in the post-test was strong (Cohen's $d = 1.99$) for the intervention group and also large (Cohen's $d = 1.09$) for the peer-assessment group. Although students from the peer-assessment group made fewer errors in the post-test, it was notable that their percentage of error-free clauses hardly changed.

Table 5. 18 Descriptive Statistics and Results of Paired Samples *t*-tests of Writing Accuracy in the Pre- and Post-tests in Two Groups (Intervention and Comparison)

Accuracy measures	Group	Pre-test vs. Post-test		
		t	*p*	Cohen's *d*
Error Free Clauses	SA	-6.64	< .001	-0.92
	PA	.30	.38	
Error/100	SA	14.23	< .001	1.99
	PA	6.98	< .001	1.09

5.4.4 Results of fluency measures

Words per minute were used to measure students' writing fluency in this study. Table 5.19 shows that students in the two groups differed only slightly in terms of fluency in the pre-test but, in the post-test, students who received self-assessment wrote significantly faster (M = 5.91, SD = .75) than those who received peer-assessment (M = 5.17, SD = 1.02) per minute, with a large effect size Cohen's $d = .99$.

Table 5. 19 Descriptive Statistics and Results of Independent Samples *t*-tests of Fluency Measures in the Pre- and Post-tests in Two Groups (Intervention and Comparison)

Fluency Measure	Group	Pre-test (T1)					Post-test (T2)				
		N	M	SD	*t*	*p*	N	M	SD	*t*	*p*
Words/Minute	SA	51	4.72	1.09	1.45	.15	51	5.91	.75	3.99	< .001
	PA	41	4.40	.96	1.47	.14	41	5.17	1.02	3.86	< .001

Based on the paired samples *t*-test results (Table 5.20), when considered within the group, writing fluency for both groups witnessed a reasonable growth from the pre- to post-test (see Figure 5.3 for an overview).

Figure 5. 3. Results of Writing Fluency in the Pre-and Post-tests

Table 5. 20 Descriptive Statistics and Results of Paired Samples *t*-tests of Fluency Measures in the Pre- and Post-tests in Two Groups (Intervention and Comparison)

Fluency Measure	Group	Pre-test (T1)			Post-test (T2)			T1 vs. T2		
		N	M	SD	N	M	SD	*t*	*p*	Cohen's *d*
Words/Minute	SA	51	4.72	1.09	51	5.91	.75	-7.15	< .001	1.00
	PA	41	4.40	.96	41	5.17	1.02	-4.12	< .001	.64

The self-assessment group outperformed the peer-assessment group. Analyses show that the self-assessment intervention, had a large effect on students' writing fluency (Cohen's *d* = 1.00), and the comparison group (peer-assessment) registered a medium effect (Cohen's *d* = .64).

5.4.5 Summary

To summarise, in this section, the most noticeable findings were the great improvement students achieved in their writing performance for both groups in the post-test. The intervention group significantly outperformed the comparison group in most dimensions, except language resources, in which both groups performed equally in both the pre-test and the post-test. Both self-assessment and peer-assessment mainly affected complexity, accuracy and fluency within-group differences, rather than between-group differences. For example, the number of clauses per T-unit in students' writing (syntactic complexity) decreased in the post-test for both groups, with the effect sizes (Cohen's d) small (.32) for the intervention group and medium (.58) for the comparison group. However, significant within-group and between-group differences were found in errors per 100 words (accuracy measure) for the two groups. Specifically, the intervention group made considerable progress, whereas the comparison group experienced no change from the pre-test to the post-test. Lastly, there was a noticeable increase in students' fluency for the two groups. However, the gain for the intervention group produced a larger effect size (Cohen's $d = 1.00$) than for the comparison group (Cohen's $d = .64$), in relation to fluency.

It seemed that students' writing performance in different dimensions was in line with their writing development in complexity, accuracy and fluency. For example, students' writing in both groups witnessed no growth in either the dimension of language resource or the measures of complexity. Similarly, with better performance in mechanics in the post-test, students indeed showed improvement in part of their accuracy measures. Moreover, students' improved fluency might contribute to the progress they had made in task achievement.

5.5 Relations between Students' Self-efficacy for Self-assessment of Writing (SESAEFLW) and Writing Performance

This section addresses the research question, "How does students' self-reported efficacy for self-assessment of writing correlate with their writing performance". To answer this question, the Pearson correlation coefficient (r) was employed as an indicator of the relationship between self-efficacy for self-assessment of writing and writing performance. The magnitude of the of a Pearson correlation coefficient r is considered using the metric whereby $r = .1 - .3$ is a small or low correlation; $r > .3 -.5$ is a medium or moderate correlation and $r \geq .5$ represents a high or strong correlation (Cohen, 1988, 1992). Links of self-efficacy for self-assessment of writing and writing performance (in five individual dimensions) for two groups were investigated in the following.

5.5.1 Intervention group: Correlation between self-efficacy for self-assessment of writing (SESAW) and writing performance

Table 5.21 shows that for students in the intervention group in the pre-test, both self-efficacy variables registered significant moderate to high correlations with all the five writing dimensions (from $r = .32$ to $r = .52$). The highest correlation ($r = .52$) was between the writing skills self-assessment efficacy and the overall writing performance, and the lowest ($r = .32$) was between writing tasks self-assessment efficacy and mechanics. Generally, students' self-efficacy for writing skills self-assessment displayed slightly higher correlations with all five writing dimensions than self-efficacy for writing tasks self-assessment. Although such a trend was apparent at the post-test as well, there was a decrease in correlations among both self-efficacy variables and all the five writing dimensions, with the exception of writing skills self-assessment efficacy and language resources which were strongly correlated with a large correlation, $r = .59$.

Of the five writing dimensions, only task achievement ($r = .28$) and language resources ($r = .29$) maintained small correlations with writing tasks self-assessment efficacy in the post-test while, for other writing dimensions, the correlations with writing tasks self-assessment efficacy were no longer significant. Apart from coherence and cohesion ($r = .22$), writing skills self-assessment efficacy still presented medium to large correlations (e.g., language resources $r = .59$) with other writing dimensions.

Table 5. 21 Correlation of SESAW and Writing Performance in the Pre- and Post-test (N = 51)

Pre-test					
Self-efficacy Variables	Task Achievement	Coherence & Cohesion	Language Resources	Mechanics	Overall
Writing tasks self-assessment	.39**	.37**	.38**	.32*	.48**
Writing skills self-assessment	.46**	.39**	.38*	.35*	.52**
Post-test					
Self-efficacy Variables	Task Achievement	Coherence & Cohesion	Language Resources	Mechanics	Overall
Writing tasks self-assessment	.28*	.13	.29*	.19	.27
Writing skills self-assessment	.30*	.22	.59**	.34*	.44**

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

5.5.2 Comparison group: Correlation between self-efficacy for self-assessment of writing (SESAW) and writing performance

With the comparison group, as shown in Table 5.22, students' self-efficacy for writing tasks self-assessment had only a low correlation and no statistical significance, with the writing dimension of mechanics in the pre-test ($r = .23$). Unlike writing tasks self-assessment efficacy, writing skills self-assessment efficacy maintained small

correlations with all the five writing dimensions; none of the effect sizes reached statistical significance.

Table 5. 22 Correlation of SESAW and Writing performance in the Pre- and Post-test (N =41)

Self-efficacy Variables	Task Achievement	Coherence & Cohesion	Language Resources	Mechanics	Overall
Pre-test					
Writing tasks self-assessment	-.02	-.09	.09	.23	.06
Writing skills self-assessment	.18	.13	.16	.23	.24
Post-test					
Writing tasks self-assessment	.04	-.06	.16	.33[*]	.18
Writing skills self-assessment	.22	-.02	.27	.10	.03

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

In the post-test, there was a significant medium correlation between writing tasks self-assessment efficacy and mechanics ($r = .33$), and small, non-significant correlations between writing tasks self-assessment efficacy and language resources as well as the overall performance. Self-efficacy for writing skills self-assessment correlated only with task achievement, language resources and mechanics weakly ($r = .22/.27/.10$, respectively), with little correlation with either coherence and cohesion or overall performance.

5.5.3 Summary

This section reported that, in the intervention group, students' self-efficacy for self-assessment of writing correlated with their writing performance to a moderate to a high degree. In the post-test, all the correlations weakened but to differing extents, with nearly

half of the correlations no longer statistically significant. For the comparison group, students' self-efficacy for self-assessment of writing did not correlate with any of the five writing dimensions in the pre-test, although a significant medium correlation was detected between writing tasks self-assessment efficacy and mechanics in the post-test.

5.6 Chapter Summary

To conclude this chapter, whereas the two groups presented notable differences in the post-test, in the pre-test they were equivalent in self-efficacy for self-assessment of writing, writing performance, and rating accuracy in self-assessment of writing.

With the correlation between students' and raters' ratings, the significant connection appears to be with students' overall writing performance for both groups in the pre-test, and task achievement for the self-assessment group in the post-test. The correlation between students' self-efficacy for self-assessment of writing and writing performance, was moderate to high at the beginning for the intervention group, yet it was weaker in the post-test, no longer demonstrating statistical significance. For the comparison group, although students' self-efficacy for self-assessment of writing did not correlate with any of the dimensions indicating writing performance in the pre-test, significant medium correlations were detected between writing tasks self-assessment efficacy and mechanics in the post-test. Qualitative findings in the next chapter supplement the results presented in this chapter.

CHAPTER SIX

QUALITATIVE FINDINGS

6.1 Overview

This section reports the views of the six students, in the intervention group, concerning self-assessment of writing from different perspectives. Data sources included semi-structured interviews, learning journals, raters' and students' comments on the writing strengths and weaknesses of their pre- and post-writing tasks, and my classroom observations. Findings of changes in students' perceptions of their self-efficacy in self-assessment of writing and writing performance, and their perceptions of, and attitude to, self-assessment of writing, and the writing rubric are presented first. I further triangulated these data with the quantitative results from the research questionnaire, to consider the research questions as stated below. It was expected that the qualitative data would provide a lens to understand better the students' development of self-assessment during the four-month intervention.

Then, the findings are reported by themes that emerged from the data to describe factors influencing students' self-assessment of writing to afford a descriptive and analytical response to the following research questions:

1) What are students' (e.g., basic/intermediate/advanced English proficiency) perceptions of, and attitudes to, both self-assessment of writing, and the rubric used for self-assessment of writing?

2) What factors influence EFL students' self-assessment of writing practices?

6.1.1 Selection of participants

Participants in this stage of the research had a range of English proficiency levels (basic/intermediate/advanced), using purposeful sampling (Mia and Yvonne were advanced; Luke and Joe were intermediate; Zoey and Grace were basic), to ensure a range of views was captured. It was also to test the assumption that English proficiency levels might affect how an individual responds to an activity such as self-assessment. Participants' detailed background information can be found in Section 4.5.2.

6.2 Self-efficacy for Self-assessment of Writing

The six students reported medium self-efficacy scores for self-assessment of writing at the beginning of the intervention with a substantial increase in self-efficacy scores by the end of the semester. An increase in confidence was evident in the students' growing ability to identify their strengths and weaknesses in writing skills and performance. Excerpts from the students' interviews and learning journals exemplifying the reasons for such an increase in self-efficacy are presented in the sections below. In the following sections, two students with similar English proficiency levels are reported together, and my classroom observations help elucidate the students' actual situation.

6.2.1 Mia's and Yvonne's views

My classroom observation showed Mia and Yvonne to be very quiet, calmly completing every self-assessment task assigned by the teacher without asking for further instructions. I was not able to ascertain how confident they were in terms of self-assessment of writing until the first interview. In this interview, both Mia and Yvonne indicated that they were moderately confident, because they knew, if only vaguely, what their weaknesses were in writing, but not their strengths, as Mia indicated in her first interview:

I would say my confidence to perform self-assessment of writing is on a mid-level. Firstly, I have never been asked this question by anyone before and I just assume I may not be too bad at it. And secondly, though I kind of know my weaknesses, I cannot tell what my strengths are. My strengths, according to my secondary teachers, should never be the focus of my learning (Mia, Interview 1).

Yvonne agreed with Mia's view, adding that, in her first interview she was moderately self-efficacious because she did not know what could be treated as self-assessment of writing as no one had explained the concept to her. By the end of the semester, both Mia and Yvonne articulated that, after engaging in self-assessment over the semester, they were fully confident in assessing both their own writing skills and writing tasks. Yvonne claimed in her second interview:

I am fully confident that I can self-assess my writing now and I am not going to be modest about that. I am familiar with the self-assessment procedure now and I can detect my grammar mistakes or unreasonable logic in my writing. Most importantly, I realised I can rely on myself to assess my writing with all the external resources without my teacher's presence. That helped me to generate effective feedback for myself, and I can understand my teacher's feedback better. My learning therefore progresses in a positive circle (Yvonne, Interview 2).

Mia also said she felt self-reliant, and more certain, when assessing her writing skills or predicting her performance in writing tasks, which she attributed to the constant setting of concrete goals in the English writing course as a part of self-assessment. By completing those goals, Mia had built up her confidence in assessing her own writing, as revealed in her last learning journal:

Before engaging in self-assessment of writing, I did not know how to describe my writing to others. But now, by practicing self-assessment regularly, I have an all-round deeper understanding of my writing. I can rely on myself to check reference books and online resources to find the good points and the weaknesses in my writing. I have set achievable goals in every stage of my writing and through assessing and achieving those goals, I now clearly know the next steps I should take to improve my writing. In the past, the teacher assessed my writing for me, and I just looked at the comments, forgot all of that, and I would make the same mistakes again next time. After self-assessing my writing, I become mindful and no longer make the same mistakes. I am filled with confidence that I can assess my writing (Mia, learning journal 3).

6.2.2 Luke's and Joe's views

Before engaging in self-assessment in the writing class, Luke and Joe were not positive about their ability to self-assess their writing. They both indicated that a lack of relevant previous experience made them feel only moderately confident about self-assessment of writing. In contrast to Mia and Yvonne's situation, Luke and Joe had a clear idea of their strengths but did not know their weaknesses. In the following excerpt, Luke pointed out his reason for not feeling confident concerning self-assessment of writing after an unexpectedly long time of thinking.

There is no false modesty. I have never tried anything like this (performing self-assessment of writing) before, or to be exact, never thought about this. And from my previous experiences, I feel my teacher knows my writing better, not myself. I do not need to think about my writing, as I would just receive my teacher's help, which cannot go wrong, right? I know I can spell most of the words right and I can

147

express my ideas clearly, but I am not able to tell my weaknesses in writing. Though my teacher did mention the areas I could improve for my writing, those words never went deep to my heart (Luke, Interview 1).

Joe similarly confessed, in the first interview, that he lacked confidence in self-assessing his writing because he did not know how to self-assess writing, as it had not been taught at all during his former education. Later, in their learning journals, however, they both demonstrated greater confidence in assessing their writing. For example, Joe wrote in his second journal:

I know the self-assessment procedure now and I can somehow assess my writing process and my writing performance. I am more confident about what I need to do before writing, during writing, or after writing. Though self-assessment involves many detailed steps, I believe I can do that little by little (Joe, Journal 2).

Luke also mentioned his growth of confidence in self-assessment of writing during the second interview; he said that he felt much more confident in assessing his English writing because he had been practicing it in different forms over the semester. He drew an analogy between his writing and his hobby, which is painting, in the following excerpt,

I love painting because I have learned painting skills from my teacher, and I feel confident to apply those skills to completing every new painting on my own ever since. It is the same as the self-assessment of writing. I feel I get better at either assessing my writing skills or writing performance because I had opportunities to evaluate my strengths and weaknesses in English writing. If someone asked me about that now, I could tell them immediately. I have learned the different

components of self-assessment through weekly practices in this semester, and then I have just become so confident in that (Luke, Interview 2).

This excerpt suggests that regular practices of self-assessment of writing helped Luke develop more confidence in his self-efficacy. Joe, likewise, explained in his third journal, that, although he is not fully confident in assessing his writing, he has gained confidence in self-assessment by constantly analysing, revising and reflecting on his writing, particularly with the help of a writing rubric.

6.2.3 Zoey's and Grace's views

Sharing the same moderate level of self-efficacy with other students at the beginning, Zoey and Grace's confidence in their self-assessment of writing came, largely, from their trust of their teacher's scaffolding. As Zoey reflected in her first interview, "I believe I am confident to perform self-assessment of writing because my teacher's instructions will guide me along the way (Zoey, Interview 1)". Grace agreed with Zoey in her first interview, stressing that although she had no previous relevant experience, she had faith in herself to follow the teacher's instruction to self-assess her writing effectively. As Grace commented:

> I think I am semi-confident concerning self-assessing my writing, because I know I could do it with the help of my teacher's professional knowledge and guidance. I should not be too bad at it (self-assessment of writing) as long as I recognise what my teacher would like me to do (Grace, Interview 1).

While Zoey and Grace felt rather ineffective when assessing their writing during the intervention, because they were unable to fully implement their teacher's instruction,

they reported salient changes in their self-efficacy for self-assessment of writing by the end of the semester. For example, Zoey explained in her second interview:

> I feel that I become more confident in assessing my English writing, from writing skills to the entire writing performance. During the writing course, I had more opportunities to practice what the teacher taught me about assessing my writing and I have tried to give myself some feedback revising my draft rather than left it there as I did prior to the intervention. I realised it should be myself that driving my study forward rather than my teacher (Zoey, Interview 2)

Grace, likewise, claimed that her efficacy for self-assessment of writing had improved greatly over the semester, as she indicated in her second interview:

> Before, when I said I am moderately confident about self-assessment of writing, in essence, my confidence came from my faith in my teacher's teaching. But now I am much confident than that time because I know such confidence has been built up in myself through regular practices. Although I think I still need at least half a year to be fully confident at it, I am sure I can assess my writing well if I carry on practicing self-assessment.

6.3 Students' Writing Performance and Rating Accuracy

The quantitative results suggested that, for students in the intervention group, there was a noticeable growth in their writing performance, which were explained to some extent by their responses in the interviews and learning journals. Students' rating accuracy in this study, refers not only to the closeness between the students' and the raters' rating, but also the similarity of the students' and raters' comments on the students' writing strengths and weaknesses. When commenting on students' writing strengths and

weaknesses, raters and students were asked to indicate, first, a writing dimension (e.g., task achievement), and then for each student's writing, why it was a strong or weak point. Although there were low correlations between the students' rating and the raters' rating, increasing similarity between the students' and the raters' comments from pre- to post-writing tasks suggests self-assessment of writing had a positive effect on students' rating accuracy. Six students' and the raters' comments in the pre-and post-writing tasks are presented in the following table for comparison.

As shown in Table 6.1, in the pre-test, there are considerable discrepancies between students' and the raters' comments, suggesting they had different ideas of a student's strengths and weaknesses in writing. In the post-test, however, the students' and the raters' comments overlapped in both the designated rubric dimensions as well as the comments on those dimensions. Students' comments are more expressive and reflective in the post-test than in the pre-test. Six students' comments on their writing performance are detailed in the following.

Table 6. 1 Raters' Comments and Students' Self-comments of Writing Strengths and Weaknesses in the Pre- and Post-tests.

Comment Sources	Strengths		Weaknesses	
	Pre-comments	Post-comments	Pre-comments	Post-comments
Mia	Language resources: I used complicated sentence structures.	Task achievement: I have answered the topic questions clearly with good examples in each part. I also have a clear structure with signal words linking my ideas.	Mechanics: Poor spelling	Language resources: I am not certain if I have used enough authentic expressions and complex structures in the right way in my writing. I need to work on various sentence structures.
Teachers	Task achievement: answered most task questions with clear opinions and examples.	Task achievement and language resources: addressed all task questions with relevant examples and good logic. Also used sophisticated expressions and attributive clauses in the right way.	Coherence and cohesion: lack of linking words to link ideas smoothly	Language resources: run-on sentences in complex sentence structures
Yvonne	Task achievement: I have abundant ideas for the topic?	Task achievement and language resources: I answered all the task questions with proper examples to support my arguments. I tried to use a range of academic expressions.	Language resources: not good at using complex vocabularies	Mechanics: I made spelling mistakes because I focused more on the words and sentence structures I used, and I neglect word-spelling. I need to make time to double check spelling alone
Teachers	Coherence and cohesion/ language resources: the writing is a pleasure to read with ideas clearly and logically presented. Also used a variety of sentence structures and sophisticated vocabulary.	Task achievement and language resources: paraphrased and addressed all the task questions with detailed examples and used sophisticated vocabularies and structures.	Mechanics: violated many grammar rules.	Mechanics: numerous spelling errors
Luke	Coherence and cohesion: Logical thoughts	Coherence and cohesion: My writing has topic sentences for each part, but no conclusion. But the whole writing is easy to follow with enough examples.	Mechanics: punctuation mistakes	Mechanics: I realise I always use tenses wrongly. I guess I don't fully understand the occasions to use the present tense. My handwriting is not easy to follow too, and I should start from improving my handwriting

Teachers	Language resources: generally correct use of complex sentence structures.	Coherence and cohesion: ideas are presented clearly and logically with good use of linking words.	Task achievement: unable to answer the task questions, extremely below the word limit.	Mechanics: many mistakes in grammar especially the use of tense. Poor handwriting disturbs reading.
Joe	Mechanics: Very few spelling mistakes.	Language resources: To answer the task questions, I have tried to use different kinds of words in different forms to achieve vocabulary variety.	Mechanics: I always spell words wrong.	Coherence and cohesion: My ideas in the body part were not logical and coherent for the topic. I need to make stronger links among my ideas and examples using transitional words.
Teachers	Task achievement: clearly answered task questions with enough examples.	Language resources: used a range of words in the right way to address task questions.	Language resources: major problems in using the right word/phrases collocation.	Coherence and cohesion: repetitive and illogical ideas, not easy to follow.
Zoey	I don't know	Task achievement: because I have a clear opinion, and I can write effective sentences to describe my examples and to support my ideas.	Task achievement: not sure about my choice	Mechanics: I always forget words' spelling and grammar rules during writing. I need more practices in grammar such as subject and verb agreement
Teachers	Language resources: used a range of topic-related words	Task achievement: most task questions are addressed with supporting details.	Mechanics: violated many basic grammar rules.	Mechanics: full of spelling errors and grammar mistakes.
Grace	Task achievement: I do not know why	Task achievement: I have used four examples to express myself clearly around the main topic. I think I have provided enough support for my argument.	No idea	Language resources: I have used simple sentence structures and basic vocabulary too often, like "think, like, etc.". I need to learn to use more advanced words.
Teachers	Coherence and cohesion: opinions are clear in general	Task achievement: most task questions are answered with relevant examples	Task achievement: half of the task questions are not addressed	Language resources: mainly used basic words and sentence structures with phrase-collocation errors.

6.3.1 Mia's and Yvonne's views

Before the intervention started, Mia and Yvonne said they believed their English writing was not good, even though they often received high marks for writing from their teachers. They were constantly frustrated that they were unable to describe how good their writing, as they indicated in the first interview:

> I always get higher grades for my writing even though I think there are so many things to be improved in my writing. The dispiriting thing is I do not know what I should work on to improve my writing. Getting higher grades means little to me and I want to know how well I actually perform in my writing (Mia, Interview 1).

> All of my friends say to me that my writing is better than theirs, but it does not mean my writing is good enough to me. I can fully express myself in my writing but I am not good at explaining my writing in detail to someone else. I am not confident enough in my writing performance, I guess that is because I do not know my writing in a deeper sense (Yvonne, Interview 1).

Although, in the pre-test, Mia and Yvonne were able to articulate the strengths and weaknesses of their writing in either a sentence or a phrase, the dimensions they identified differed markedly from those of recognised by the raters. For instance, whereas Mia identified her strength as in the dimension of language resources, in which she used many complex sentence structures, the raters suggested her strength was task achievement, with task questions answered clearly using examples. Similarly, Yvonne and the raters had different opinions as to her strengths and weaknesses at the beginning. For example, Yvonne briefly indicated that her strengths were presenting enough ideas

for the topic but the raters believed she performed better in terms of coherence and cohesion as well as language resources.

By the end of the semester, however, Mia and Yvonne scored themselves similarly to the raters' rating and referred to their confidence when talking about their writing performance. As they explained,

> Engaging in self-assessment in-and-out of the classroom provided me more opportunities to interact with my writing. I have never tried to have conversations with either my writing or myself before. I think my writing performance has been greatly improved as I can improve my writing with my own feedback (Mia, Interview 2).

> I believe my writing performance is much better than before because it generally satisfies me now. I know my writing in a deeper way by assessing its structure and language and, at the same time, referring to many external resources helps me to conduct effective self-assessment (Yvonne, Interview 2).

As well as from their self-reported enhanced efficacy in writing performance, Mia and Yvonne, in the post-test, were more accurate in assessing their writing strengths and weaknesses and their comments indicated greater congruence with the raters' comments. For example, although Mia and the raters still did not agree on her weaknesses, half of their comments on her strengths were the same; while Yvonne and the raters identified the same dimensions for her strengths (task achievement and language resources) and weaknesses (mechanics) in the post-test.

6.3.2 Luke's and Joe's views

In the pre-test, Luke and Joe were somewhat confident about their writing performance. Both the raters and Luke and Joe suggested a medium range for their writing performance based on the quantitative results. In the first interview, however, they suggested it was largely because they always receive mid-range scores for their writing, and they were just used to rating themselves in a similar way.

> I believe my confidence in writing performance is on a medium level. That confidence does not come from me, but from the medium grades my teacher usually assigned to me. It is not necessary for me to know my writing very well as long as my teacher always makes the judgement (Luke, Interview 1).

> I am moderately confident in my writing performance as, previously, my teacher said my writing is OK. I have never thought or asked my teacher about what does OK mean. I do not know how to interpret that comment in detail in my writing (Joe, Interview 2).

As with Mia and Yvonne in the pre-test, although Luke and Joe were able to articulate their strengths and weaknesses, their comments on writing did not match those given by the raters. In the post-test, however, their comments were closer to the raters' comments. For instance, Luke chose the dimension of coherence and cohesion as his strength and mechanics as his weakness from the pre- to the post-test. in the post-test, and the raters agreed with him in terms of his strengths and weaknesses. Later in the second interview, Luke mentioned the reasons for choosing his writing strengths and weaknesses.

> When I put down my strengths and weaknesses in the pre-writing test, I did not take it seriously. I just randomly picked two dimensions that I believed I could be

good at or weak in and wrote them down. But performing self-assessment during this semester has prompted me to think about my writing seriously. I have studied coherence and cohesion in detail, and I know how coherent my writing is; I also realised that my terrible handwriting seriously affects reading and understanding. Now, I am very confident in my writing performance generally (Luke, Interview 2).

Joe also said he felt more confident about his writing performance after engaging in self-assessment for a semester. In the post-test, the comments by Joe, and the raters' comments on Joe's writing, indicated his progress in writing. Unlike his comments at the beginning of the semester, Joe reported greater inner motivation as he had committed more time and effort through self-assessing, which had built his confidence in writing in the post-test. As he stated in his third learning journal:

I definitely feel more confident than before about my writing performance, and I believe my writing is getting much better. My writing is chaotic before, but after doing self-assessment regularly I feel there is something that I can control in my writing. For example, I can tell some of its problems and make improvements accordingly. I did spend much time finding a better way to assess my writing; I hoped my efforts would pay off.

6.3.3 Zoey's and Grace's views

In the pre-test, Zoey and Grace were reluctant to talk about their confidence in writing performance because that question required too much thinking by them. When asked, Zoey replied in her first interview:

Why bother to think about my writing or my confidence? I believe the easiest and safest way is always to apply the doctrine of the mean, which means I am neither very confident nor not confident, and my writing is neither excellent nor the worst. In that way, I feel safe and better among my classmates (Zoey, Interview 1).

Grace said she kept thinking and searching for the right words to describe her confidence in her writing performance, but in the end, gave up saying that:

Honestly, that question never occurred to me before and I do not know the answer because I do not know my writing at all. I only write when I am asked to, and I stop thinking about it once I finished my writing, or maybe I do not even think about my writing when I write it. Writing has always been a one-off activity for me (Grace, Interview 1).

Zoey's and Grace were not able to respond to their strengths and weaknesses in writing either. For example, they could only comment such as, "I don't know" or "No idea", or merely indicated a specific dimension, which shared no similarity with that pointed out by the raters, without further comments in the pre-test. Nonetheless, later in the post-test, it seemed that Zoey's and Grace's ability to identify strengths and weaknesses in their writing had been enhanced over the period of a semester. They were not only able to identify their strengths and weaknesses in the same dimension as the raters, but were also capable of adding some more details to express themselves (see Table 6.1 for details). In the following excerpt, Zoey explains the reasons for her huge boost of confidence.

I think practicing self-assessment of writing regularly has helped me to build up great confidence in my writing performance. I feel happy now because for the first time, I know what my writing strengths and weaknesses are, which are not told by my teacher but found out by myself. During this semester, the teacher either assured my assessment or enlightened me about the next steps. I feel I have more ownership in my writing (Zoey, Interview 2).

Zoey expressed a greater willingness to use self-assessment at different stages of writing and to exert more effort in learning to write while actively controlling her writing process. All these appeared to work together, contributing to her confidence in writing performance. Grace expressed similar growing confidence in her writing performance, as she mentions in the following.

I cannot say if my writing has become different but for sure, I know I am much more confident in my writing performance than before. I come to know my writing, especially my strengths and weaknesses through self-assessing it regularly with the help of the rubric. Self-assessment also expands my horizon and prompts me to reach out for other resources to improve my writing (Grace, Interview 2)

6.4 Students' Perceptions of, and Attitudes to, Self-assessment of Writing and the Writing Rubric

In the first two months of the writing class, the six students pointed out that self-assessment of writing and using a rubric independently were entirely novel for them. They were not used to assessing themselves because they were accustomed to being assessed by the teacher. The students also admitted to a long-standing perplexity about self-assessment, how to self-assess, and how to utilise a rubric in their writing. Students

reported that their initial impression of self-assessment of writing was it was pointless and a waste of time. It was only after the mid-term that students started to become accustomed to self-assessment of writing and were more receptive to its meaning and practices. It appears that students began to have a deeper and broader understanding of self-assessment of writing.

Although students displayed marked differences in their perceptions of, and attitudes to self-assessment, students' responses to the interviews and in their journals suggested that, by the end of the semester, all students found self-assessment of writing useful in multiple ways. Five out of the six students agreed that they enjoyed self-assessing their writing because it gave them a feeling of self-control in their learning. They enjoyed the sense of ownership in self-assessment; it appeared to encourage them to be actively involved in improving their writing and to devote time to assess their writing for other subjects after class as well.

All the students pointed out that the most difficult aspects of self-assessment of writing were grading themselves and improving their essays to a satisfactory level after self-assessment. They were challenged also by how to revise their writing to a satisfactory level after self-assessment, as the rubric did not give them formative feedback about either their success in self-assessing writing or about what to do to remediate issues in their writing. While students believed frequent one-to-one meetings with their teacher would enhance their progress in self-assessment of writing, they were aware that it was not feasible for the teacher, due to the large number of students, as exemplified in students' responses in the following sections.

6.4.1 Mia's and Yvonne's stories

Mia and Yvonne, who had relatively higher English proficiency, did not experience any frustration when self-assessing their writing over the semester, although, in the beginning, they judged self-assessment of writing as not useful. For instance, Yvonne wrote in her first learning journal:

> I cannot see the point of performing self-assessment since I have already tried my best when I wrote my essay. I spend a lot of time rethinking everything again and again with barely no improvement. I do not know if I have self-assessed my writing in the way that my teacher expected, because since I started schooling, no one has ever taught me about this.

The comment, "I have no idea if I perform self-assessment in the way that my teacher expected" was echoed by Mia who struggled with whether her approaches to assess her writing work were professional. Although Mia agreed that self-assessment of writing is a way to identify one's strengths and weaknesses, she thought constant self-assessment was time-consuming. Later, they agreed that self-assessment of writing is a novel but rewarding learning experience through which to explore their capability in writing. As Mia commented in her second interview:

> Self-assessment for me is no longer just about assessing my writing but more about motivating myself to critically reflect on learning and other real-life conditions. I have control in this journey to explore my writing competence, which I find can not only be applied in the current course, but also in other courses and similar situations in the future.

Yvonne agreed with Mia, adding that self-assessment of writing is not only for self-exploration, but also for self-development in learning skills. They admitted there were difficult times when they were not able to make the desired progress in writing through self-assessment. However, Yvonne considered such times as valuable, as it motivated her to engage in deeper self-assessment, that is, she claimed she was not motivated by the grades but focused on what she had learned about her writing during the assessment process. For example, Yvonne explained in her second interview that, "self-assessment is indeed challenging, and I might have given up without my teacher's encouragement. However, self-assessment is about learning, it is not a competition. I have become more conscious about what I know and what I do not know, and I think this is beneficial for the rest of my time in college and my future career".

The writing rubric, for all students at the beginning of the intervention, was unfamiliar and rather overwhelming. Yvonne described it as follows,

> When I got the rubric, I was like Wow, am I able to use this rubric? I meant it is very exhaustive. There is a huge amount of information. I have never seen a rubric like this before. To read the rubric thoroughly and to apply the rubric to my writing is not very hard, but I mainly use the rubric as a reminder for my structure check after writing, not relying on it much (Yvonne, Interview 1).

Mia similarly commented on her use of the rubric in the first interview, "I only refer to the rubric when I finish my writing. I am afraid that consulting the rubric before writing will constrain my mind. For me, the rubric is just a reference for format checking (e.g., font and layout)". Both Mia and Yvonne pointed out that when they applied the rubric to their writing at the beginning of the course, they placed their writing at the mid-level of the rubric because they did not want to appear arrogant about their writing to others.

Despite knowing their writing was better than average, Mia and Yvonne believed it showcased their modesty when others, including their teacher, checked their self-assessment outcome. They said they felt safer not being seen as successful in the presence of others. By the end of the semester, however, they rated their writing at the top level confidently, realising that being honest to themselves is more constructive to improve their writing.

6.4.2 Luke's and Joe's stories

Unlike Mia and Yvonne, Luke and Joe were frustrated by the concept of self-assessment of writing at the beginning. Initially, they had trouble understanding the meaning of, and the reasons for, performing self-assessment of writing. As Luke stated in his first interviews, "self-assessment is only a quick fix for handwriting, spelling, and simple grammar mistakes. It is just a mere formality, and there is no point to take it so seriously since it is all about my ideas". Likewise, Joe commented in his first journal, saying that "self-assessment is pointless. I cannot see any differences in my writing so far and I feel uncomfortable doing this because I do not trust myself". Joe added in his interview later:

> I have not been taught anything like self-assessment before since primary schooling. I do not know what that means. Usually, my teacher is the authority, and I have no say in my learning. But, suddenly I was told I could actually assess my writing. The feeling is so unreal. I ask myself, am I eligible to assess myself? The answer is always negative for me and I think it is a waste of time (Joe, Interview 1).

Not only were Luke and Joe uncertain about their ability to assess their own writing, they were also unsure about the definition of self-assessment of writing. For example, Luke noted in his first interview after pondering for a long time:

> Self-assessing my writing truly challenged my mind. It has been just so hard, and I do not know how to describe that feeling. Honestly, I am not sure what self-assessment of writing is. For now, I just copied my writing to another piece of paper again to make sure my handwriting is neat enough. That is also a kind of self-assessment, right? (Luke, Interview 1)

By the end of the semester Luke and Joe started to articulate a clearer understanding about self-assessment of writing, as seen in Luke's second interview:

> Self-assessment is a complicated and systematic process that happens before, during and after writing. It is way more than a quick check. I can self-assess my writing sentence by sentence, paragraph by paragraph or as a whole. It also requires a lot of critical thinking to figure out my strengths and weaknesses in different stages (Luke, Interview 2).

Joe's entry in his last learning journal provides a good illustration of the change in students' view of the self-assessment of writing. He wrote, "self-assessment is more than a task or simply assigning a grade for my writing. It is constant retrospection and reflection that helps me to recognise my actual writing ability".

Further, Luke and Joe commented that it was time-consuming to understand the writing rubric as for them it was rather wordy and obscure; therefore, initially, they were unwilling to refer to the rubric when self-assessing their writing because they were unable to respond after referring to it. Later, however, students requested a more detailed

rubric, which suggests further actions but maintains conciseness. During the second half of the semester, Luke and Joe started to accept the guiding function of the rubric in self-assessing their writing; they said they now viewed the rubric as if they were the examiners and as a reminder, to improve their writing either in word choices or essay structure. As Joe explained in the second interview:

> After using the rubric for around three months, the rubric was already somehow printed on my mind. I think about the rubric dimensions unconsciously while writing. And I tried to make my writing better to match a higher level of the rubric. Though the whole rubric thing is still somehow vague to me, it is very helpful to structure my thoughts by providing a reference. By referring to the rubric, my writing is now controllable to some extent (Joe, Interview 2).

The following excerpt from the second interview also exemplifies the change in Luke's understanding of the writing rubric.

> At first, I was reluctant to refer to the rubric even if my teacher asked me to. Because the rubric looked so overwhelming. But later I realised the more I referred to the rubric, the clearer my writing became, and the more confident I became, feeling like I was a professional examiner. I still do not know how to improve my writing according to my own feedback after assessing it, but at least now I know what I need to improve to make my writing better. That is a big step for me (Luke, Interview 2).

6.4.3 Zoey's and Grace's stories

The classroom observation recorded that Zoey and Grace rarely answered the teacher's questions even when they were asked, and constantly fell asleep during the class. Grace expressed her feelings in the first interview,

> I am not focused on self-assessment. I just passively do whatever my teacher asked me to. Because according to my previous experience, the teacher used to override all my thoughts about learning. I have never self-assessed before, and now by no means would I self-assess my writing more than once. I just grade myself roughly and that is all. Because however well I do, it is always the teacher's call in the end. I do not want to waste my time if my effort is not valued and recorded (Grace, Interview 1).

Zoey exhibited similar indifference to self-assessment of writing practices in her class, saying in her first interview, "I do not care about self-assessment. To me, it is only another formality. I just followed my teacher's instruction and soon I get bored because I cannot see the point of self-assessing my writing. It is my teacher's job to grade my writing, not mine". For students such as Zoey and Grace, self-assessment of writing, in the early stages of the course, was more likely to be a task to fulfil the teacher's requirements rather than something they would initiate.

Close to the end of the semester, while Zoey and Grace began to appreciate the value and process of self-assessment, they also admitted they took 'a surface approach' to self-assessment of writing and how self-assessing their writing plunged them into self-doubt, Zoey wrote in her third journal:

I finally started to enjoy assessing my writing, not grading my writing, I mean I enjoy the process. I know I am still not good at it because all I can do is to correct a few spelling or grammar errors, and I am not capable of replacing some words for better expressions or restructuring my essay after self-assessment. But I have learned how much I can rely on my own efforts as the teacher's help is not always available. I feel…very hopeful, you know, I realised this is my study, and I should take responsibility for it. But I also feel doubtful about my overall English writing or learning capability and whether I could help myself to improve my writing (Zoey, Journal 3).

Whereas Grace also started to realise the significance of self-assessment, she did not experience satisfaction from self-assessing her writing, because her limited progress in writing constantly frustrated her. For instance, in the second interview, she noted that:

This whole self-assessment thing is still so beyond me. On the one hand, I enjoy the process as I get to know my writing better. On the other hand, I feel terrible as I barely have any progress in assessing my writing after investing so much time and effort into it. But that is not the point. It is so hard to give myself feedback in English and I have no idea how to revise my writing to a satisfactory level with my own feedback. I began to self-doubt all aspects of my essay. The process of self-assessing my writing is like, keeping reminding myself how terrible my English is, and ironically, what I can do is only to give myself a better mark. I know I do not deserve that, but at least, it makes me feel better (Grace, Interview 2).

Zoey and Grace devoted most of their time to translate the writing rubric from English to their mother tongue (Mandarin) at the beginning because it was difficult to understand.

They said they believed the rubric is of great importance in self-assessment of writing as they were required to utilise it all the time. When Zoey and Grace could not follow what was taught in the writing course, it was observed that they switched their attention to translating the rubric during the class. Zoey indicated in her first interview that, "I spent days looking up the unfamiliar words in the writing rubric. I have translated the whole rubric into Mandarin and finally it looked much clearer and understandable to me". Likewise, the following comment from Grace in her first interview illustrates her feelings facing and using the rubric.

> I do not know what to do with the rubric other than translating the whole thing into Mandarin, so it is easier for me to use it later. It is the first time that I view my writing in such a detailed way. I feel overwhelmed and challenged but at the same time excited to learn more about the rubric. The writing rubric seems helpful to orient myself to my writing (Grace, Interview 2).

Zoey and Grace were more skilled at applying the rubric to their writing near the end of the semester. They said they appreciated the usefulness of the rubric as it helped them to foster critical thinking not only in English writing, but also in other aspects of language learning, as well as in real-life situations. Zoey's comment in her learning journal illustrated such points:

> I feel I am much more critical than before after using the rubric to assess my writing. Because the rubric has different aspects and I started to view my learning in other subjects from various perspectives too. Self-assessing my writing has opened a window of critical thinking for me (Zoey, Learning journal 2).

Grace also claimed she was more critical after using the rubric to assess her writing for a semester. She expressed the opinion, however, that her self-assessment was still rather limited as she did not understand the rubric fully and needed more time to digest the rubric contents.

6.5 Influential Factors in Students' Self-assessment of Writing Practices

Through inductive thematic analysis, data collected from interviews, learning journals and classroom observation, three elements were revealed that were likely to influence students' practices in self-assessment of EFL writing: English proficiency level, secondary teacher feedback practices, and cultural norms concerning self-esteem and modesty. Students' views representing these themes are reported in the following sections.

6.5.1 English proficiency level

The students' English proficiency levels, particularly low English proficiency, seemed to affect their understanding and practices of self-assessment as well as how they used the self-assessment of writing rubric. For example, Zoe and Grace reported that they used Mandarin, their native language, predominantly when self-assessing their writing. They reported that for them Mandarin is the better medium for self-assessment practices and self-representation in the writing class. Zoey explained her reasons in the first interview:

> It's difficult for me to express myself clearly in English, not to mention using those English grammar jargon like prepositions and conjunctions. But I would understand everything about self-assessment if the teacher had used Mandarin in her teaching. I spent days translating the English version of the self-assessment

writing rubric completely to Mandarin and the Mandarin rubric is much friendly for me when assessing my writing (Zoey, Interview 1).

During the classroom observation, it was noted that Zoey and Grace constantly fell asleep. Later they attributed their in-class sleepiness to the fact that they could not concentrate when their teacher taught self-assessment procedures in English and they felt overwhelmed learning a new concept in a foreign language. They pointed out that because of their low language competence at the beginning of the self-assessment course, they were only able to detect spelling errors. By the end of the semester, they felt that their self-assessment of writing was still superficial, and they were unable to consider the essay organisation and logic. Zoe and Grace believed that their low English proficiency level prevented them from enjoying or making progress in self-assessment of writing.

In contrast to Zoey and Grace, Mia and Yvonne, with high English proficiency, mentioned in both pre- and post-interviews that they mostly focused on the appropriateness of logic and ideas in their self-assessment of writing. It appeared easier for Mia and Yvonne to understand the self-assessment procedure and the rubric, and they rarely experienced frustration (see more details in section 6.2.1 and 6.3.1). For instance, Mia commented in her first learning journal:

Understanding the rubric is not very hard for me; I have been trying to make sure my self-assessment of writing goes beyond spelling or simple grammar checking. I am keen to improve my logic and to support my thoughts with proper examples. I need to check more references and I know I could improve my writing in this way (Mia, Learning Journal 1).

Luke and Joe, in the middle of the English proficiency range, could understand the writing rubric without translating it into Mandarin. As for Zoey and Grace, however, they said that they felt their limited English prevented them from taking further actions to improve their writing. For example, Luke, in his second interview, said that he was not satisfied with simply picking out spelling or grammar mistakes in self-assessment, and while wanting to enhance his writing coherence and cohesion, realised he needed a better grasp of English sentence structures to express himself.

6.5.2 Secondary teacher feedback

The feedback practices of their former teachers were another salient factor identified as influencing students' self-assessment of writing practices. In the first interview when asked how they self-assessed their writing, all students commented that their secondary teachers' feedback practices as influencing their self-assessment of writing. For example, Luke responded that:

> To be honest, I do not know how to assess my writing at the beginning. But I clearly remember that my secondary English teacher always corrected every mistake in my writing with a red ink pen and she mainly gave me feedback on my handwriting. She said good handwriting is powerful enough to hide other flaws in one's writing and it was the easiest thing I could do to improve my writing. So, I have decided to follow what I was taught before; that is why I have paid a lot of attention to improve my handwriting when assessing my writing (Luke, Interview 2).

Similarly, Yvonne and Zoey mentioned how their confidence in self-assessment and their self-assessment of writing practices had been deeply affected by their secondary English teacher. For instance,

> For passing exams' sake, all my secondary English teachers kept asking us to check spelling mistakes or simple grammatical errors. That was deeply rooted in my mind as a standard for good writing. Therefore, when I was first asked to self-assess my writing, I just checked my grammar and spellings repeatedly as those were emphasised by my secondary teachers, and I, myself, was not trusted to assess other dimensions of my writing by that time (Yvonne, Interview 1).

> I am not confident in assessing my own writing because I am no good at English writing according to my secondary teachers' feedback. How can I self-assess my writing in university then? I was told that my writing could only get better if I closely followed the teacher's feedback. So, there is actually no 'me' in my writing, it always about what my teachers want. I feel lost in self-assessment because I do not know if I am doing it to the teacher's expectation (Zoey, Interview 1).

The long-lasting effect of their own teachers' feedback practices on students' self-assessment of writing was also recorded in Grace's second learning journal. For instance, in the following excerpt, from the second learning journal entry, Grace commented on how she felt about self-assessing her writing saying:

> I have been struggling when I self-assess my own writing as my secondary teacher and my current teacher keep fighting in my mind; My secondary teacher usually gives most students, including me, a medium mark, as I guess that avoided hurting some students' feelings if his or her work was not good enough. So, I have become

accustomed to grading myself in the mid-range even though I know I am not good

enough to reach that range. But my current teacher tells me that honesty matters

in our self-assessment. It is hard to shift my mind concerning the way I grade

myself in self-assessment in a short time (Grace, Learning Journal 2).

6.5.3 Cultural norms: Modesty and self-esteem

Cultural factors such as self-esteem and modesty also seemed to influence students' self-assessment of writing practices. When the six students were asked how they used the writing rubric during self-assessment, regardless of their English proficiency levels, all indicated that they would give their writing three out of five based on the criteria, that is, the mid-level. They preferred not to position themselves too high or too low according to the writing rubric, even though their self-assessment of the writing process was not seen by others, In doing so, students believed they demonstrated both Chinese virtues of modesty and saving face. The following excerpts from Mia and Grace, with contrasting English proficiency levels, Mia high and Grace low, are examples.

> I am sure my writing is better than the average, but I always rate my writing
>
> starting from the mid-level of the rubric. It never occurs to me I could first match
>
> my writing from the top level of the rubric because I have always been taught that
>
> thinking highly of yourself is no good for your study or your future. How dare I
>
> rate my writing from the top level? Even if I am excellent, it is for others to judge,
>
> not myself. This is the value that I grow up with, modesty (Mia, Interview 1).

> I know I am no good in English writing, but I will by no means assess my writing
>
> from the bottom level of the writing rubric. I need to save face in this, just for
>
> myself. I know if I rate myself from the mid-level, I would still have to move

downwards later. But that short illusion of not losing my face, and I am actually

not that bad <as it> helps me maintain or gain a little confidence in performing

self-assessment of writing (Grace, Interview 1)

Students' concern about losing face is also apparent in their unwillingness to seek the

teacher's assistance during self-assessment. Surprisingly, such concern is evident with

all students, as the following interview excerpts from Luke and Zoey exemplify:

> My English teacher is the last person I will turn to for help. I distance myself from
>
> teachers since primary school because they always act as the authorities. I do not
>
> want to be seen as needy for my teacher's help. As a grown-up, if I ever need help,
>
> I will turn to my male friends, not even females. If they cannot help me, I would
>
> rather leave the questions there, unless my teacher senses my concerns and reaches
>
> me first (Luke, Interview 2).

> I feel ashamed of asking my English teacher to help me with self-assessment of
>
> writing because I know I am not smart and this will be a big favour. I am afraid I
>
> will ask silly questions and my teacher may ridicule or think less of me (Zoey,
>
> Interview 2).

Even students with high English proficiency such as Mia and Yvonne tended to keep all

their concerns about self-assessment of writing to themselves rather than sharing them

with their teacher. They believed that asking for help was incompatible with the image

of excellent students that they had tried to create in their teacher's mind, as exemplified

in Mia's and Yvonne's learning journal responses. Mia wrote, "my teacher thinks I am

smart and independent, and she speaks highly of me. So, I do not want to show her my

weaknesses. I want to rely on myself and other resources to improve my writing (Learning Journal 2)". Yvonne, similarly, replied that:

> Revealing my weaknesses to my teacher will somehow make her think less of me. I will only reach the teacher when a problem has been bothering me for ages and I cannot find an answer. Otherwise, after practicing self-assessment of writing during this semester, I have enough faith in myself to solve most of the problems in my study (Learning journal 3).

6.6 Chapter Summary

This chapter takes a closer look at six students' engagement in self-assessment of writing over the period of a semester time. The findings reveal an enhanced self-efficacy for self-assessment of writing and writing performance for all students. The students' increased accuracy in rating was reflected, not only in the similarity between their own comments and the raters' comments, but also in comments expressed by students. The above findings also described the students' reactions, adaptation, and perception of self-assessment of writing and the writing rubric. Students' self-assessment of writing seemed to be largely influenced by their own language proficiency and the feedback process of their former teachers, as well as cultural norms.

CHAPTER SEVEN

GENERAL DISCUSSION

7.1 Chapter Overview

As reported in Chapter Five, results from the questionnaire and the writing tasks indicated that the intervention group's use of self-assessment had a positive effect on Chinese EFL learners' self-efficacy for self-assessment of writing, writing performance, and rating accuracy. For the intervention group, the quantitative results also demonstrated that their self-efficacy for self-assessment of writing was significantly positively correlated to their writing performance in the pre-test, but not in the post-test. In Chapter Six, the qualitative findings from students' interviews and learning journals supported the favourable effects of using self-assessment in the EFL writing classroom and revealed the benefits of self-assessment intervention for students' perception of self-assessment and its relevance to their future learning. Additionally, the qualitative findings revealed that students' self-assessment of writing practices was not only influenced by their English proficiency, but also by their own secondary teachers' feedback practices. Chinese culture norms, such as modesty and self-esteem, seem to play an important role in students' self-assessment of writing practices as well.

In this chapter, the findings reported in Chapter Five and Six are presented and discussed in relation to the research questions. Each section starts with a brief summary of the quantitative results and qualitative findings to the specific research question. These are then discussed with reference to the current study context, the wider context of the theoretical framework and relevant studies on self-assessment.

7.2 RQ1a: What Effects Does the Use of Self-assessment Have on Students' Self-efficacy for Self-assessment of EFL Writing?

Descriptive statistics indicated that students from both the intervention group and the comparison group, who possessed a moderate level of self-efficacy for self-assessment of writing at the beginning (see section 5.2.2 for details), experienced significant growth in the two dimensions of self-efficacy for self-assessment of writing after one semester (see section 5.3.1 for details). These results had some consistency with Bandura's (1986) contention that students' confidence to accomplish the writing tasks should increase after a writing intervention. The increased self-efficacy, documented for the comparison group, also corroborates earlier findings that peer-assessment contributes to learners' self-assessment development (Bouziane & Zyad, 2018).

However, students from the intervention group had a larger effect size (Cohen's $d = 3.74$ and 2) than the comparison group (Cohen's $d = 1.56$ and 1.42), which suggested comparatively greater improvement in the two dimensions of self-efficacy for self-assessment of writing for the intervention group. The self-efficacy increase for the intervention group might be explained by the qualitative findings, which recorded that six students from the intervention group, with different English proficiency levels, explained why such an increase occurred. These explanations are summarised as four aspects: Direct and regular experiences of self-assessment of writing practices; established familiarity with self-assessment procedures over the semester; the revival of personal agency and self-motivation in the learning process; and the lecturer's scaffolding, in-time feedback and encouragement. Students, it appears, were more efficacious in assessing their work after engaging in self-assessment. They realised they

should be driving their study and not their teachers; and therefore, they committed more time and effort to self-assessment practices (Bai et al., 2019; Ross et al., 1999; Xu, 2019).

These findings are broadly consistent with previous research into the sources of writing self-efficacy, one of which, individuals' accumulated mastery experience, has been identified as the most influential in enhancing self-efficacy (Bandura, 1997; Bruning & Kauffman, 2015; Butler, 2018; Schunk & Pajares, 2010; Usher & Pajares, 2008). In the current study, students' previous moderate levels of self-efficacy for self-assessment of writing could be explained by their lack of opportunities to practice; later, the positive development in students' self-efficacy, is consistent with the view that direct tasks and practices involving self-assessment reinforce students' self-efficacy level accordingly (Alishah & Dolmaci, 2013; Andrade et al., 2009; Black & Wiliam, 1998; Dörnyei, 2001; McMillan & Hearn, 2008; Ross, 2005; Schunk, 2003; Van Reybroeck, Penneman, Vidick, & Galand, 2017). Moreover, in this study, the positive social persuasion, for instance, verbal or written encouragement and support from the teacher, served as an important influence that motivated students carrying on self-assessing their writing and further improved students' self-efficacy for self-assessment of writing (Bandura, 1997; Bai et al., 2019; Ruegg, 2018; Sun & Wang, 2020).

Further, the quantitative findings of this study are congruent with the limited previous studies measuring Chinese non-English major students' writing self-efficacy, which indicated that Chinese English major students expressed a medium level of self-efficacy in the self-assessment of the writing domain (Sun & Wang, 2020; Tang & Xu, 2011; Woodrow, 2011). The students' moderate level of self-efficacy for self-assessment of writing may have been due to their lack of experiences in class; activities involving students in the assessing or learning process are considered time-consuming (Lam,

2016), in the traditional Chinese classroom environment in which teaching and learning of EFL writing are driven mostly by exams (Hui et al., 2020; Lee, 2014). Students' lack of self-knowledge or self-awareness in their writing may be another explanation for the medium self-efficacy for self-assessment of writing students reported in this study. Students, it would seem, did not have a clear understanding of their learning progress, or their writing strengths and weaknesses (Hobson, 1996), possibly because Chinese university students are accustomed to being reactive to their teachers' proactive management of their writing (Liu, 2002).

Previous studies have measured students' writing self-efficacy for writing skills and tasks (Pajares, 2003; Shell et al., 1989). The present findings add a different dimension to the existing literature by determining writing self-efficacy from the perspective of self-assessment with EFL tertiary students. Unlike Woodrow's (2011) study reporting that Chinese college students have less confidence to perform writing skills than writing tasks, in this study students from both groups had relatively higher self-efficacy with writing skills self-assessment than with writing tasks self-assessment (see section 5.3.1 for details). It is possible that in the Chinese context grammar-learning is not only the focus in class, but also the emphasis of exercises and examinations in different forms (Zheng & Adamson, 2003). Therefore, mastering writing skills may attract students' attention more than completing writing tasks. Students may also possess comparatively lower self-efficacy in assessing their writing tasks since they are not aware that they have the agency to assess, or they have rarely been offered the opportunity to practice self-assessment (Cassidy, 2007).

Contrary to Shell et al.'s (1989) findings that high achievers hold high self-efficacy, the present results suggest that EFL students' English proficiency levels may not be

congruent with students' self-efficacy levels. Further examination of the learners' perceptions of self-assessment, in this study, revealed that although the six participants in the qualitative study had varied English proficiency levels, their self-efficacy for self-assessment of writing at the beginning was similar. An explanation of these findings, indicated in students' interviews, may be their lack of knowledge and experience with self-assessment of writing when reporting on their ability to carry out such activities, and hence they were cautious about their capability to self-assess their writing (see section 6.2 for details).

The six students from the qualitative study clearly harboured doubts about their capability to self-assess their writing, and that influenced their perceived self-efficacy initially. By the end of the intervention, however, as in the case of Grace, even though she pointed out that her limited progress in self-assessment of writing constantly frustrated her, she was cognisant of the increase in her self-efficacy for self-assessment of writing. Such results are in line with Bandura's (1986) notion that even though learners judge their progress at a lower level, it may not lead to lower self-efficacy.

Even though students from the comparison group received peer-, not self-assessment, during the intervention period, their self-efficacy for self-assessment of writing also increased significantly in the post-test; it is possible that peer-assessment fosters students' capability to conduct effective self-assessment (Black et al., 2004) which then heightens their self-efficacy for self-assessment. As this study adopted a relatively new instrument (the researcher-developed *Self-efficacy for Self-assessment of EFL Writing* questionnaire) to measure Chinese students' self-efficacy for self-assessment of writing, more studies are needed to provide further evidence of its utility and validity.

7.3 RQ1b: What Effects Does the Use of Self-assessment Have on Students' Writing Performance as Well as Their Linguistic Improvements in Terms of Complexity, Accuracy, and Fluency?

7.3.1 Writing performance

The findings show that students' writing performance improved in both intervention and comparison groups. The increase in the intervention group's (self-assessment) writing performance had a greater effect size (Cohen's $d = 1.89$) than that of the comparison group (peer-assessment, Cohen's $d = .78$). The intervention group also performed better than the comparison group in the post-test in the writing dimensions of task achievement, coherence and cohesion, and mechanics (see section 5.4.1 for details). The results show some consistency with previous studies in which the positive effects of self-assessment on overall writing quality improvement have been demonstrated (e.g., Andrade & Boulay, 2003; Andrade et al., 2008; Andrade et al., 2010; Birjandi & Siyyari, 2010; Lam, 2013; Mazloomi & Khabiri, 2018; Oscarson, 2009). The qualitative findings complemented the quantitative findings with all six students indicating they felt their writing had improved through practicing self-assessment regularly which had helped them have a deeper understanding of their writing; they took the initiative to use external sources to improve their writing.

Both groups exhibited within-group significant growth (Cohen's $d = 1.06$ and 1.03 for the intervention group and comparison group respectively) but no between-group differences in the post-test in language resources, the only dimension in which this was evident. A possible explanation might be that during the pedagogical (self-assessment) intervention, the teacher's instruction did not emphasise students' use of vocabulary and

sophisticated sentence structures, therefore, self-assessment and peer-assessment influenced students' use of language resources to a similar extent.

7.3.2 Linguistic developments: Complexity, accuracy, and fluency

Apart from exploring the effects of self-assessment on students' writing holistically, this study has extended previous studies (e.g., Birjandi & Hadidi Tamjid, 2012; Birjandi & Siyyari, 2010; Oscarson, 2009) by considering writing linguistic developmental indices, namely, complexity, accuracy, and fluency. These findings, described in the following sections, may be unprecedented because the effect of self-assessment on the development of writing linguistic features has not been the focus of sustained attention and only little empirical work exists in this area.

7.3.2.1 Complexity

Writing complexity was considered from two perspectives: syntactic and lexical complexity. The two groups had no between-group changes in students' syntactic complexity (mean length of *T*-units and number of clauses per *T*-unit) and lexical complexity (mean segmental type-token ratio and lexical density) in the post-test (see section 5.4.2 for details). Although students' writing performance improved, it was not necessarily more complex. This result, however, was similar to Ross et al. (1999), which reported that using self-assessment in students' narrative writing does not discernibly improve linguistic structure. Previous findings, in contrast, had also suggested that students who received self-assessment treatment outperformed the peer-assessment group in writing complexity (Bouziane & Zyad, 2018).

It is possible that students' lack of improvement in writing complexity after the intervention (see Table 5.14 for details) was due to the relatively short intervention

duration (four months). Previous research has suggested that for L2 writing at the university level, approximately one-year instruction is needed at least to show substantial changes in linguistic complexity (Ortega, 2003), or to see any effect of a specific educational treatment (Wolfe-Quintero et al., 1998). The Chinese students in the present study may have needed more time to become accustomed to self-assessment of writing and then to show greater variation and sophistication in their use of the English language.

Although there were no between-group differences, in the post-test there was a significant (Cohen's $d = .57$) increase in the lexical density for the intervention group. As Read (2010) argued that higher lexical density implies higher literacy, therefore, the self-assessment intervention may have somewhat improved students' literacy, which is students' ability to write. Contrary to expectations, this study found a significant decrease in one of the complexity measures, the number of clauses per T-unit for both groups (Cohen's $d = .32$ and .58 for the intervention group and comparison group respectively). Such results support the contention that students' language development does not follow a linear path, and that different linguistic dimensions in writing may not develop simultaneously (Sadler, 1989).

It can be concluded that engaging in self-assessment of writing seems to have little effect on students' writing complexity indices because the complexity measures could be sensitive to the experimental condition (self-assessment). Another possibility is that the task selection and design of this study might be cognitively demanding for students to develop further writing complexity as students were unlikely to respond under pressure, especially when self-assessment was not familiar to them (Casal & Lee, 2019; Foster & Skehan, 1996).

7.3.2.2 Accuracy

Similar to previous studies (Bouziane & Zyad, 2018; Skehan & Foster, 1997), the intervention group outperformed the comparison group significantly in general accuracy indices (error-free clauses and errors per 100 words), and within-group gains were much larger than the between-group gains (see section 5.4.3 for details). These results extend and complement Oscarson's (2009) study confirming the positive effects of self-assessment on students' writing accuracy.

Greater accuracy after engaging in self-assessment of writing may be the result of accuracy being of great concern as students control language forms of writing, and of Chinese students attaching more importance to language forms than other aspects of writing (Ding & Zhao, 2019). Furthermore, since the general accuracy indices cover a broad range of error types (Wolfe-Quintero et al., 1998), it may be easier for students to detect and revise their errors in writing.

According to the students' interviews, self-assessment may have heightened their sense of control in the use of language and writing behaviours. Moreover, during the intervention, students from the case studies indicated that they engaged in mostly surface self-assessment, so that their initial focus in self-assessment was the correction of errors. Students' attention in error-correction may stem from the Chinese writing teachers' traditional approach in responding to students' submitted writing by correcting every error with red ink (Lee & Coniam, 2013). Students' attention to grammar and other forms of language use would naturally result, with, a greater level of accuracy likely to be attained (Skehan & Foster, 1997; Van Reybroeck et al., 2017).

7.3.2.3 Fluency

Consistent with previous literature (Bouziane & Zyad, 2018), the intervention group outperformed the comparison group significantly in the fluency measure (words per minute), but the within-group gains were much larger than the between-group gains (see section 5.4.4 for details). The observed increase in students' speed of production could be the result of the extra planning time through pre-writing self-assessment (e.g., self-reflection question: Do I have a clear outline for this essay?), which enhanced students' confidence in the writing topic and formulation of writing content. These factors could assist students in the intervention group to write significantly faster than their counterparts in the comparison group (Abdel Latif, 2013; Kellogg, 1990). The results of this study also support evidence from previous studies (e.g. Van Reybroeck et al., 2017), showing that a self-assessment intervention helped to improve students' text production.

In summary, results from the present study imply that self-assessment benefits students' language development to some extent. There are trade-off effects, however, among complexity, accuracy and fluency as shown in previous empirical evidence. It appears that due to students' limited attention and capability, as well as possible anxiety in facing a novel assessment form (Nawas, 2020), it is difficult to focus on the three aspects simultaneously, and so students prioritise the development of one in their writing (Skehan, 2009; Skehan & Foster, 1997). In this study, the intervention group's gains in writing accuracy and fluency seemed to be at the expense of complexity. Further studies are needed to develop more effective self-assessment practices and instructions to achieve balanced and long-term progress in the areas of complexity, accuracy and fluency, without development in one area compromising the other areas.

7.4 RQ1c: What Effects Does the Use of Self-assessment Have on Students'

Rating Accuracy as Compared to the Raters' Assessment?

Results indicated that there were only moderately significant correlations of the students' and the raters' ratings for both groups' overall writing performance, with very little change from pre-test to the post-test in the correlations. The non-significant correlations for coherence and cohesion, language resources, and mechanics reduced further in the post-test. Similar results were reported by previous studies (Leach, 2012; Longhurst & Norton, 1997; J. Zhang, 2016), suggesting that when self-assessment practices are learning-oriented, students judged their overall writing similarly to external raters, but there were discrepancies in rating for the individual dimensions. While there was a significant correlation ($r = .38$) between the students' and the raters' rating in the dimension of task achievement, in the other individual writing dimensions, the correlation between the students' and the raters' rating decreased in the post-test. Such results are congruent with Liu (2002), which showed Chinese students' ratings correlated with the raters' only in task completion due to their limited English proficiency.

Further, results of this study differ from those in previous research that reported self-assessment training could improve students' self-rating reliability (Blanche & Merino, 1989; Oscarson, 1989), and that students' peer-assessment scores were more accurate than self-assessment scores when compared to the external rating (González-Betancor et al., 2019). The present study concurs broadly with Birjandi and Siyyari (2010) that neither self-assessment nor peer-assessment enhanced EFL English-major undergraduate students' rating accuracy significantly. Furthermore, based on the relatively poor correspondence between the students' rating and the raters' rating in the

post-test (see section 5.3.2 for details), in this study, it appears that self-assessment did not enable learners to make judgements about their writing performance, especially for individual writing dimensions, comparable to those made by the raters.

There are several plausible explanations for the reduced correlation between the students' rating and the raters' rating for individual writing dimensions (except for task achievement). Firstly, it may be the effect of engaging in self-assessment, in which students report negatively on themselves when reflecting their self-assessment. This can lead students to self-doubt and unnecessary anxiety (Nawas, 2020; Pajares & Urdan, 2005; Yan 2020), and negatively affect judgement on their work. Secondly, alternatively as indicated in previous literature, students with low-level proficiency (Harris & Brown, 2018; Liu, 2002), may overestimate themselves in self-assessment because of ego protection or lack of relevant knowledge (Earl & Katz, 2006; Matsuno, 2009; Pintrich, 2002; Sullivan & Hall, 1997; Topping, 2003). In the present study, for instance, Grace, a student with basic English proficiency level, indicated in her second interview that self-assessment made her realise how terrible her English was and she started to doubt every aspect of her writing; and to make herself feel better, she gave herself a better grade.

Moreover, as learners (novice raters) were probably given insufficient classroom instructions and time to digest the self-assessment or peer-assessment materials and practices, they were likely less prepared to assess their work, such as vocabulary and sentence structures in writing, using the provided detailed rubric, than the experienced raters (Xu, 2019; J.Zhang, 2016). Another reason may be that the students had inadequate teacher feedback during the intervention for accurate student self-assessment

development. At this early stage, learners may have still lacked understanding and experience in self-assessment (González-Betancor et al., 2019; Oscarson, 2009).

However, one exception was that, for the intervention group, the students' and the raters' ratings for the dimension of task achievement correlated significantly in the post-test (from $r = .19$ to $r = .38$). The result aligned partially in alignment with Liu (2002), in which students' rating and teachers' ratings correlated significantly for writing task attainment, but not for the use of language resources. This result also suggests that the self-assessment intervention may improve students' rating accuracy in general writing perspectives such as task achievement over a semester time; however, for students to develop better judgement of their linguistic performance, they need extra time and practice.

From a formative perspective, students' rating accuracy might have improved after self-assessment as students brought their comments on writing strengths and weaknesses more in line with the raters (González-Betancor et al., 2019). Those findings further support the statement that "self-assessment may not lead to an improved essay; however, it may lead to an enhanced insight into the strengths and weaknesses of the essay the students will submit" (Hanrahan & Isaacs, 2001, p. 65). The fact that students were able to generate more accurate, perceptive and constructive comments for their writing strengths and weaknesses in the post-test supports the argument that self-assessment can assist students to refine their judgements, to gauge the quality of their work better, and to nurture and identify their writing potential (Panadero & Jonsson, 2013). For example, as shown in Table 6.1, in the post-test, students with varied English proficiency levels were more aware of their writing strengths and weaknesses, similar to the raters, and students reflected on how to improve their weaknesses.

The nature of students' post-test comments seemed to become more positive and perceptive than their pre-comments according to Hattie and Timperley's (2007) four levels of feedback focus, including the level of task performance, task processing, self-regulation, and the personal level. Taking Zoey (basic English proficiency) as an example, in the pre-test, she commented simply "I don't know" for her writing strengths; in the post-test, however, she picked the dimension of task achievement, further adding "because I have a clear opinion, and I can write effective sentences to describe my examples and to support my ideas". Such comments not only reflect Zoey's enhanced confidence in her self-assessment, but also indicate Zoey's enhanced self-regulatory proficiencies in her writing practices.

These results seem to be consistent with Chen's (2008) study, in which students' comments, after engaging in assessment practices for two cycles, were comparable to the teachers, and showed a raised awareness of their strengths and weaknesses. Such change may have resulted from the teacher's instruction and encouragement during self-assessment practices, students' gradual internalisation of the teacher's commentaries and from learning to appreciate their own work. Students' interviews and learning journals also suggested an enhanced competence in judging their writing strengths and weaknesses in the post-test (see section 6.3 for details).

Although students in this study seem unable to assess their work reliably in relation to the raters' scores, it is still necessary to include self-rating during self-assessment. Learning to self-rate not only helps students reflect on their work, but may also help teachers reduce their marking time and to free time for other classroom activities (Boud, 1989; González-Betancor et al., 2019). Therefore, the question is not about how accurate students can be in rating their writing, but about how to establish students' confidence

in assessing their performance and to foster a deeper understanding of writing through self-rating. Using self-assessment formatively, arguably, could be one approach to realise those goals (Liu & Xu, 2017).

7.5 RQ2: How does Self-efficacy for Self-assessment of Writing (SESAW) Correlate with Students' Writing Performance?

Whereas the intervention and the comparison group had comparable self-efficacy levels and writing performance in the pre-test, the correlations between SESAW and writing performance varied for each group. For instance, the intervention group demonstrated in the pre-test, moderate to high correlations (from $r = .32$ to $r = .52$) between the two SESAW variables (writing skills and writing tasks) and students' writing performance (five dimensions). For this group, the correlation of self-efficacy for writing skills self-assessment and writing performance was higher than that of self-efficacy for writing tasks self-assessment (see section 5.5.1 for details). The results corroborated the findings of other pertinent studies in this area suggesting there are significant correlations between writing self-efficacy and writing performance as well as writing self-efficacy being predictive of writing performance (Bruning et al., 2013; McCarthy et al., 1985; Pajares, 2003; Pajares & Johnson, 1994; Shell et al., 1989). For the comparison group, however, only small correlations with no statistical significance were evident between SESAW and writing performance. While it is difficult to explain this result, it may be that students' individual differences, and teacher differences within those two groups, are likely to impact students' learning or rating outcomes (Nielsen, 2019).

Further, in the post-test, correlations between self-efficacy and writing performance dimensions dropped to different degrees for the intervention group with the exception of the correlation between self-efficacy for writing skills self-assessment and language

resources, which increased from $r = .38$ to $r = .59$. The relatively weak associations were unanticipated and apparently have not been found elsewhere. The decrease may signify inconsistencies in students' self-efficacy beliefs and their actual writing performance during self-assessment (Schunk & DiBenedetto, 2016; Schunk & Usher, 2011). In other words, the development of students' writing in the different dimensions did not consistently parallel the development of self-efficacy; this may be true especially for difficult and novel tasks such as self-assessment in the EFL context. For instance, as emerged in students' interviews, although learners with basic English proficiency were highly efficacious about self-assessing writing after the self-assessment intervention, they admitted they felt demotivated and they still could not write in English to the desired level because of their lack of knowledge in regulating their learning resources. These findings are contrary to previous studies which have suggested that learners' strong sense of confidence helps improve writing performance and stimulates greater interest and effort in writing (Pajares, 2003; Schunk & Pajares, 2010).

Unlike Woodrow's (2011) study with Chinese students which reported that self-efficacy was a potent predictor of writing performance, the associations between self-efficacy and writing performance gradually diminished in the present study. In the pre-test, the participants most likely did not have a comprehensive understanding of what comprises self-assessment of writing when they reported their judgment on their ability to conduct such tasks.

Regarding the comparison group, only self-efficacy for writing tasks self-assessment correlated with mechanics moderately ($r = .33$), with no discernible or obvious relationships revealed between self-efficacy and other writing dimensions (see section 5.5.2 for details). One possible reason for this lack of relationship could be that the

participants might judge themselves capable of or inefficacious in self-assessing their writing, but then they received peer-assessment during the semester for their writing instead. That may result in poor correspondence between self-efficacy for self-assessment of writing and writing performance.

From the results in this study, given the inconsistency of self-efficacy and writing performance, self-efficacy for self-assessment of writing does not appear to be a major factor predicting Chinese students' writing performance. It is important, therefore, for future research to include other aspects of writing self-efficacy to understand better their links to students' writing performance.

7.6 RQ3 What Are Students' Perceptions of, and Attitudes to, Self-assessment of Writing, and the Rubric Used for Self-assessment of Writing?

7.6.1 Perceptions of and attitudes to self-assessment of writing

Consistent with the literature, at the beginning of the intervention, participants had difficulties in understanding the concept of self-assessment and positioning the 'self' in their self-assessment (Andrade & Du, 2007). Students indicated in their first interviews that they self-assessed their writing only because their teacher asked them to, and that the self-assessment form was primarily teacher-driven (see section 6.4 for details). In contrast to previous findings of students' resistance to being involved in the self-assessment process (Sambell et al., 2013), students in this study have shown mainly passive acceptance. It may be that the participating students have an ingrained idea that the teacher is the authority with power in the classroom; and that they had to obey authority and be introspective (Littlewood, 1999; Xu, 2019).

There appeared to be a consensus among students that self-assessment is a novel but meaningless and time-consuming task; a task which was supposed to be the teachers' responsibility (Harris & Brown, 2013; Lee, 2011a; Tan, 2009). A possible explanation is students' lack of confidence in self-assessing their writing; initially, students' self-efficacy for self-assessment was only medium. The level of self-efficacy is salient for students to initiate and sustain efforts in their learning, and that might explain why students showed limited interest in self-assessment of writing (Schunk & Usher, 2011; Zimmerman, 2011). Another reason for students' negative attitudes towards self-assessment could be that they lack the knowledge of self-assessment. They reported initially that they had never been taught about self-assessment and did not know what exactly self-assessment was (see examples in section 6.4.2). Their responses about their attitudes were predictable, as for students, especially in the Chinese education context, self-assessment is rarely listed in the Chinese teaching syllabuses and materials; students are strongly performance-oriented, with grades the impetus (Liu & Yu, 2017). Lack of domain knowledge, however, would inevitably impede students' development in effective self-assessment of writing.

Since self-assessment does not instantly contribute to students' grades (Harris & Brown, 2018), this study identified a challenge in addressing students' anxiety that the time they invest on self-assessment of writing is not proportional to the perceived short-term benefits (Hanrahan & Isaacs, 2001; Wanner & Palmer, 2018). Students' agency in self-assessment should be encouraged and realised by teachers' intervention and support (e.g., through one-to-one meetings) in students' self-assessment in the early stages, to dispel students' overriding concerns and sense of insecurity about their ability to self-assess; (Nielsen, 2019). It is incumbent upon teachers when introducing self-assessment to

consider students' knowledge base, to raise students' awareness of the importance of self-assessment and to help students better exercise their agency (Brown & Harris, 2013).

After the intervention, students' uniform endorsement of the positive role that self-assessment plays in writing represented a change from passive reception to active involvement. That meant students became accepting of self-assessment and able to view themselves not only as learners but also as assessors (Nawas, 2020; Sadek, 2018). Students' positive attitude change may be the outcome of their enhanced self-efficacy beliefs about self-assessment of writing; individuals' self-efficacy levels may affect how they perceive and react to a task (Bandura, 1986). Such findings are consistent with previous studies reporting students' favourable attitudes towards self-assessment of writing after regular and extended practices, despite still indicating that self-assessment was difficult (Andrade & Du, 2007; Chen, 2008; Mazloomi & Khabiri, 2018; Oscarson, 2009; Sullivan & Hall, 1997).

Students' difficulties in grading themselves and lack of improvement in their writing after self-assessment seemed to lead to frustration and discouraged responses (Nawas, 2020; Pajares & Urdan, 2005; Yan 2020) from students with basic English proficiency in this study. For instance, both Zoey and Grace became self-doubting or even demotivated about every aspect of their writing and English learning as their self-assessment of writing confronted her with the reality of her level of English proficiency. Such findings suggest that self-assessment may not work for everyone (Panader, Brown, et al., 2016). For students, who are weaker academically, the effects of self-assessment may be counterproductive in the short term, and that might be the reason that other studies have posited the effectiveness of self-assessment is evident only when the intervention lasts for a long time span (Andrade et al., 2010; Wang, 2016).

As Andrade and Valtcheva (2009) claim, it is important for teachers to inform students about both the short-term effect (e.g., on a particular essay) and the long-term effect (e.g., on self-regulated learning) of engaging in self-assessment and to guide weaker students to use self-assessment formatively and not focus on self-assessment as self-grading (Brown et al., 2015).

Other students from the qualitative study realised that self-assessment was an important skill, which helped them develop critical thinking about their writing; self-assessment was also a medium through which they could take responsibility for their own learning (González-Betancor et al., 2019; Hanrahan & Isaccs, 2001; Orsmond et al, 1997; Stefani, 1994). Mia, illustrated the viewpoint by commenting in her second interview that "self-assessment for me is no longer just about assessing my writing but more about motivating myself to critically reflect on learning and other real-life situations and I have control in this journey of exploring my writing competence". Mia appeared to be aware that the capability they developed during self-assessing their work is of benefit both in the current course and for future situations (Boud & Soler, 2016; Panadero et al., 2019).

7.6.2 Perceptions of and attitudes to the rubric used for self-assessment of writing

This study affords support for the effectiveness of rubric-referenced self-assessment (Andrade & Du, 2007; Andrade et al., 2008; Andrade et al., 2010). Through self-assessing their writing against the explicit criteria provided with the rubric, students gradually internalised the criteria and became more proficient in using the writing rubric to consider their writing from different perspectives. By the end of the intervention, confident in using the criteria of the rubric, students seemed to review their learning progress not only as students, but also as capable self-assessors (Panadero et al., 2016).

Students also indicated several difficulties they encountered when they applied the rubric to their writing, which were similar to those described by Wang (2016). For example, the rubric itself was overwhelming for all students because of its length, and at first sight, its obscurity. Contrary to intermediate and advanced level students, who used the rubric as it is, students with basic English proficiency expressed their frustration in spending days to translate the rubric into Chinese for their later use. From classroom observation, I found that weak students started to use class time to translate the rubric when their teacher used English-medium instruction to teach the self-assessment of writing. It appeared that whether weaker students focused in class depended on the language the teacher used for instruction. Therefore, it is suggested that a rubric, which includes bilingual instructions, assessment materials and exemplars of various levels should be available for all students, especially the weaker ones (Sullivan & Hall, 1997; Wanner & Palmer, 2018). A further suggestion is that EFL teachers instruct students in their native language at the initial stage of self-assessment of writing to familiarise them with the relevant procedures, alternatively to ask students to assess only what they already know and can do with the English rubric to lessen their anxiety.

As reported in previous studies, students in the present qualitative study suggested a rubric with a wider scoring range to grade their writing from one to ten instead of one to five, and clear guidance for actions to act on after the self-assessment (Andrade & Du, 2007; Wang, 2016). A challenge for future researchers and practitioners is to devise a more detailed rubric for students, which retains its conciseness so that students will not find it intimidating. To address this dilemma, it is suggested that students be invited to discuss, formulate and establish their own criteria so that they understand both the rubric and their needs better.

The difficulties identified in using the rubric for students to grade themselves imply that more attention should be paid to the primary role of self-assessment in facilitating and developing students' learning repertoire rather than assigning a grade to their work. It has been suggested, therefore, that formative self-assessment practices should be included in classroom instruction (Huang, 2016), because "counting self-assessments toward final grades may turn students' attention away from the quality of their work toward getting a high grade, thereby compromising their honesty and their focus on learning" (Andrade & Du, 2007, p. 172).

7.7 RQ4 What Factors Influence EFL Students' Self-assessment of Writing Practices?

The data from the interviews with the six students suggest that their self-assessment of writing practices was possibly influenced and mediated by a number of internal and external factors (Paris & Paris, 2001). The following sections discuss three main factors.

7.7.1 English proficiency level

It is claimed that English learners' learning experiences are affected by their levels of English proficiency (Abedi, 2010; Zhang & Zhang, 2018), and that self-assessment is a cognitively demanding process for EFL writers. Previous studies (Wang, 2016; Wanner & Palmer, 2018; Yoon & Lee, 2013), however, did not fully investigate the impact of English proficiency on tertiary students' involvement in self-assessment. Findings of this study, therefore, add to the literature in suggesting that tertiary learners' variability in English proficiency could largely influence their self-efficacy, participation and experiences in self-assessment of writing, especially for basic level students. Previous research has reported that language proficiency not only affected students' self-

assessment accuracy (Blanche & Merino, 1989; Meihami & Razmjoo, 2016), but also influenced the depth of students' understanding and engagement in self-assessment of writing (Brown & Harris, 2013; Topping, 2003) or students' use of rubrics (Wang, 2016). This study, consistent with some earlier studies (Liu, 2002; Longhurst & Norton, 1997), found that basic level students overestimated their writing capability to protect their ego (see section 6.5.1 and 7.7.3 for details).

Students with basic English proficiency levels also focused more on surface features (e.g., spelling and grammar error correction) during self-assessment because they did not fully understand the marking rubric or know how to deploy the resources and feedback they achieved. As discussed in the previous section, weak students reported that the rubric used in this study was linguistically complex. For such EFL learners, this meant a dual challenge, as they needed to understand not only the foreign language itself but also the academic requirements of that language (Abedi, 2010; Nawas, 2020).

Teachers also need to provide intermediate level students with additional guidance and time to respond to their identified issues, concerned with generating and making productive use of their self-feedback after self-assessment, and to develop their capability in self-assessment (Black et al., 2004; Harris & Brown, 2018). Advanced students, however, did not experience similar difficulties when practicing self-assessment. High-level students in this study were cognisant of the areas in which they needed to improve, and were confident in addressing these matters themselves. Compared with basic level students, advanced level students experienced less apprehension during the intervention. After the intervention, however, basic and intermediate-level students were able to self-assess themselves in a more open and robust way, similar to that of the raters' assessment. Such findings concur with Yoon

and Lee's (2013) conclusion that the effects of self-assessment intervention could be greater on basic and intermediate-level learners than on high-level learners.

The findings of this study support the argument for including individuals' differences in course design and for the writing instruction to develop self-efficacious and competent writers (Zhang, 2013). Based on the findings in this study, to ensure that self-assessment benefits high-level students equally, teachers are uniquely positioned to orchestrate more effective self-assessment techniques according to advanced learners' needs. Teachers might also reflect on how they can optimally incorporate learning opportunities and activities in the writing classroom to address the discrepancies in students' English language proficiency beforehand in order to move each student to the next level of understanding self-assessment of writing.

7.7.2. Secondary teacher feedback

Similar to Yu and Hu's (2017) study, in which secondary English teacher's writing feedback practices influenced students' subsequent peer feedback, this study found that secondary writing teachers' feedback practices also had an ongoing impact on tertiary students' self-assessment of writing practices. Based on the students' report (see section 6.5.2 for detail), it was apparent that secondary teachers' feedback affected not only students' later confidence to self-assess but also students' approach to practicing self-assessment of writing.

Students' assumption that they should not be responsible for their writing resulted in insufficient confidence to self-assess their writing in this study. During their secondary schooling, teachers had not encouraged them to be confident writers during the writing process, and so they did not trust their capability in self-assessing their writing. Due to

students' disempowerment in their previous writing experiences, it was difficult for them to position the 'self' in later self-assessment practices. Students in this study commented that they carried out self-assessment mainly because their teacher asked them to; they were concerned whether their self-assessment was performed to the teacher's expectation rather than their own. The tension was created between students' desire to meet the teacher's expectations and a focus on themselves in self-assessment (Andrade & Du, 2007).

In this study, when they self-assessed their writing, students from all proficiency levels tended to focus on their handwriting, spelling and grammar errors as these had been aspects of great importance, according to their secondary teachers. For most of the students, their attention during self-assessment of writing was directed to such mechanics and therefore prevented them from engaging in deeper self-reflection. As argued by Usher and Pajares (2008), "preconceptions lead to predispositions, and much time and patience are required to overcome the predispositions toward perceived incompetence that come to imprison these young minds" (p. 790). It is not easy to alter students' entrenched beliefs fostered by their educational background, as students' former learning and assessment experiences shape how they respond to assessment later in their lives (Boud, 1995; Zhang, 2016). The secondary teachers' approach to feedback may be attributed to the Chinese educational system, in which exam-oriented and teacher-centred classroom environments reduce students' interest in, or willingness to, consciously deploy different strategies to facilitate their learning to write in English (Yang & Gao, 2013).

A way to change students' entrenched beliefs about their focus in self-assessment of writing is first, for writing teachers to shift their focus in writing away from mechanics

(e.g., spelling, grammar errors and handwriting). Then, writing teachers need to explore the nature and purposes of self-assessment of writing with students, encouraging students to place equal importance on other aspects of writing (e.g., organisation and coherence) to improve their self-feedback practices. Teachers could provide, thus, the basis for preparing and supporting students' agency as assessors in self-assessment of writing.

The importance of preparing pre-service teachers and current secondary teachers with an understanding of assessment literacies and self-assessment is a key finding of the present study, which is supported by earlier studies (e.g., Meihami & Razmjoo, 2016). The traditional teacher-dominated pedagogical methods may discourage students from engaging in their writing; students' self-efficacy for self-assessment of writing, and students' use of self-assessment of writing practices is contingent on the feedback from their teachers (Lee, 2017). The influences of such contextual and external factors cannot be ignored.

7.7.3 Cultural norms: Modesty and self-esteem

Students' inaccurate self-assessment has often been ascribed to cultural influences, such as expectations of modesty and self-esteem; those influences were also evident in this study. Students' experiences and practices in self-assessment of writing reflected those commonly shared cultural virtues in contemporary China (Brown, 2005; Chen, 2008; Matsuno, 2009; Meihami & Razmjoo, 2016). Consistent with earlier findings that advanced students inclined to underrate themselves and weaker students tended to overrate themselves (Boud & Falchikov,1989; Brown & Harris, 2013; Dochy et al., 1999; Orsmond et al., 1997), students in this study of all proficiency levels, rated their writing in the middle range of efficacy at the beginning. These findings resonate with Harris and

Brown's (2018) contention that students reconciled their self-assessment with their perceived social standing of themselves in the classroom.

In students' interviews, the advanced-level students rated themselves in the middle-range of the rubric as they wanted to display modesty, whereas the basic-level students also matched their writing with the middle-level of the rubric because they felt better about themselves when doing so. These behaviours seemed to be culture-bound. For instance, Zoey, with basic English proficiency, indicated that she applied the doctrine of the mean when rating herself in the middle-range because her writing was not the worst in the class. The doctrine of the mean is a doctrine of Confucianism indicating that an individual should maintain balance and harmony and never act in excess. Zoey, however, misinterpreted the meaning of such doctrine and misused it in her self-assessment. Similarly, the advanced level students felt psychologically safe, too, if their self-assessment did not appear successful to others. Although, in the present study, students' self-assessment was not exposed to their classmates, students were still concerned about the outcome reflecting, perhaps, the importance of psychological safety on students' self-assessment practices as proposed in other studies (Andrade & Brown, 2016; Yan et al., 2019).

As noted above, students with basic English proficiency overrated themselves deliberately to maintain a sufficiently high level of self-esteem to enable them to continue their later study with confidence. The possible dishonesty of students in this study when grading themselves against the self-assessment rubric, suggests again, the pitfalls of using self-assessment summatively in the course (Brown et al., 2015). A solution to this problem is probably located in teachers guiding students to focus on their

academic growth rather than ego protection to achieve more accurate student self-assessment (Harris & Brown, 2018).

7.8 Chapter Summary

In summary, since, in the observed data, it was the lecturer's first time teaching self-assessment in her writing class, effective implementation of self-assessment may not be guaranteed. The discussion in this chapter has touched upon the potential effectiveness of self-assessment intervention on Chinese students' self-efficacy for self-assessment of writing, writing performance, and attitudes to self-assessment. Possible factors that influence students' self-assessment practices were also discussed. As well as students' enhanced self-efficacy levels and writing performance, the findings have demonstrated that students were also more perceptive of and reflexive about self-assessment. However, as novice self-assessors, students' self-efficacy levels (for self-assessment of writing) were not consistent with their writing development; students' rating was also inconsistent with the raters' rating.

Further empirical research needs to be conducted to provide evidence not only for the balanced progress of self-efficacy and writing performance, but also for the development of students' rating accuracy quantitatively and qualitatively. Moreover, it is evident from this discussion that self-assessment of writing is unlikely to be a 'one size fits all' practice. Rather, it is contingent on many variables, such as students' English proficiency levels, previous teachers' feedback approaches and cultural factors that have been considered in this study.

CHAPTER EIGHT

CONCLUSION

In this final chapter of the thesis, I first present a summary of the key findings of the current study, followed by the implications of this thesis for theory and pedagogy in the field of second language acquisition and learning assessment. In the remainder of this chapter, limitations and recommendations for future research are discussed in the end.

8.1 Summary of Main Research Findings

Drawing on theories of constructivism, social cognition, and formative assessment, this study employed mixed methods to collect quantitative and qualitative data with multiple instruments. The first part was the quasi-experimental study with the questionnaire, writing tests, and self-assessment tasks; the qualitative data including students' semi-structured interviews, learning journals and classroom observation formed the multiple-case study for data triangulation. Building on existing studies on self-assessment of writing (e.g., Andrade et al. 2008; Andrade et al., 2010; Liu, 2002; Liu & Brantmeier, 2019; Sullivan & Hall, 1997), this study generated results as further empirical evidence in support of the effectiveness of self-assessment on students' self-efficacy and EFL writing development in the Chinese context.

The Self-efficacy for self-assessment of EFL writing questionnaire was developed and validated for reflecting Bandura's (1986) postulation that self-efficacy belief is "a set of differentiated set of self-beliefs linked to distinct realms of functioning" rather than "an omnibus trait" (p. 36). Adopting a quantitative approach, Part One of the main study investigated changes, as seen from the pre-test to the post-test, in Chinese learners'

perceived self-efficacy beliefs in self-assessment of writing as well as their writing performance. Results showed that the self-assessment intervention was successful in helping the students in the intervention group to enhance their self-efficacy beliefs and writing performance (including linguistic development), compared with students from the comparison group, who engaged in peer-assessment. Findings also suggested that self-efficacy for self-assessment of writing might not be a reliable predictor of writing performance in the EFL context, as the relationship between the two variables could not be explained in the pre-and post-test for both the intervention and the comparison groups.

Part two of the main study described the students' experiences of self-assessment of writing. Qualitative findings revealed that the implementation of self-assessment helped students make not only a more accurate judgement of their strengths and weaknesses in EFL writing, but also generate constructive self-feedback for further action. Students' perceptions and attitudes towards self-assessment of writing were also shown to change positively over a semester. Three factors, namely, English proficiency, secondary teacher feedback, and cultural norms that may influence students' self-assessment of writing practices, were identified.

8.2 Contributions and Implications

Though the current study may not be able to make a strong claim on the effects of using self-assessment in contrast to traditional assessment practices due to the unforeseen decision on the research site to use peer-assessment instead of using the previously agreed normal teacher assessment in the comparison group, it contributes to the field of second language education from the following perspectives.

8.2.1 Theoretical contributions

This quasi-experimental project, as a whole, contributes to the development of EFL writing research by employing multiple perspectives, namely, constructivism, self-regulation and self-efficacy from socio cognitive theory as well as the formative assessment to examine how self-assessment can be learned and taught more effectively, in EFL writing and, more generally in second language education. This study has further conceptualised self-assessment and provided insights into how the three theories were reflected in the self-assessment of EFL writing practices and EFL writing in the Chinese context (Harris & Brown, 2018). The current study has not only enriched the assessment literature regarding using self-assessment in the EFL context but also expanded our understanding of the effects of self-assessment on the development of linguistic features in students' writing; it has also emphasised the significant role that the continuous interaction of behaviours, personal factors (affective) and learning environments play in students' self-assessment knowledge construction and academic performance (Bandura, 2006; Powell & Kalina, 2009). In addition, the findings confirm that self-efficacy belief is a multidimensional construct in a specific domain, that is, EFL writing in the Chinese context (Bruning et al., 2013; Chen & Zhang, 2019; Teng et al., 2018).

Moreover, evidence of the positive effect of self-assessment implementation not only supports the effectiveness of rubric-referenced self-assessment but also reflects the view of cognitive constructivism and social cognition that individuals, as capable agents, could activate, construct and adjust their knowledge, perceptions and behaviours to achieve their learning goals (Dann, 2014; Shepard, 2000; Zimmerman, 2011). From a social constructivist point of view, the self-assessment-based intervention strengthened the interaction between teachers and students as teachers' scaffolding and

encouragement not only developed student capacity in self-assessment, but also relieved students' psychological burdens when they performed self-assessment of writing in English as a foreign language (EFL) (Lee, 2017; Orscarson, 2009; Rust et al., 2005; Vygotsky, 1962). For example, in this study, all the students indicated in interviews that their teacher's encouragement and reassurance were vital for them to progress further in the self-assessment of writing and to relieve the anxiety in taking responsibility for their writing.

8.2.2 Methodological contributions

Methodologically, in this study, two contributions are made towards investigating the effects of using self-assessment on students' self-efficacy and writing improvement. First, the *Self-efficacy for Self-assessment of EFL Writing* questionnaire was successfully developed and validated based on the quantitative investigation. Using student samples from 15 Chinese universities, the questionnaire is also proposed as a potential instrument to be employed in other contexts, providing evidence and insights in students' self-efficacy in a specific domain.

Another methodological contribution lies in the quasi-experimental design of this study. Previous research has tended to adopt either a quantitative or qualitative approach for investigating the role of self-assessment in the area of writing with non-English major students. Nevertheless, using English major students as the sample, this study, which consisted of two complementary parts, used multiple research methods (quantitative and qualitative) for data collection and triangulation. Such a research design has the potential to inform, and to be used for, valid and in-depth investigations.

8.2.3 Pedagogical implications

From a practical perspective, the exploration into using self-assessment in the EFL writing classroom to understand students' self-efficacy for self-assessment of writing has some pedagogical implications. Initially, the newly developed instrument (SESAEFL questionnaire) could be used as a diagnostic tool for students to appraise their confidence in, and cultivate their awareness of, self-assessment of writing (writing skills and writing tasks). For EFL learners, the scores calculated from the instrument, provide an indication of their confidence levels in self-assessment of writing as the base for becoming become more proactive in self-assessment to improve their writing accordingly. EFL researchers and practitioners could use the SESAEFL as a precursor to empirical research, or classroom instruction, to elicit EFL learners' self-reported data to predict students' response to a self-assessment intervention in the writing classroom. Results from the SESAEFLW can also be utilized by EFL practitioners to gain insights into students' psychological conditions, such as the dimensions of writing in which students have less confidence. Therefore, teachers could provide timely encouragement and feedback to help students develop higher levels of self-efficacy beliefs accordingly (Bai et al., 2019; Kim et al., 2015). From such self-assessment data, teachers could also set teaching goals and practices that are more sensitive to students' needs, tailor instructional practices appropriately, which may lead to greater academic gains from self-assessment of writing (Schunk & Ertmer, 2000, Woodrow, 2011, Xu, 2019; Zimmerman, 2011).

Another implication concerns the empirical evidence which supports self-assessment-based intervention in EFL writing classrooms. As Chinese students are accustomed to a centralised education system and tend to accept so-called "standard" or authoritative

evaluations (Chen, 2008; Xu, 2019), the findings reported here provide information for the integration of self-assessment as a regular element into curriculum and course material design (Nielsen, 2019; Tan, 2007). For example, a top-down supportive assessment environment at a university or department level should be developed to strengthen teachers' legitimacy and capability of using self-assessment with students over a period of time (Harris & Brown, 2018; Lam, 2019). EFL teachers are advised to apply self-assessment practices in progressive stages (see section 2.2.1.1) to activate and develop students' self-assessment competence gradually (Andrade & Du, 2007; Rolheiser, 1998). Teachers also need to provide students with ongoing scaffolding and encouragement to use self-feedback in their writing, and to sustain students' confidence in self-assessing their writing and the positive effects of self-assessment. As Chinese teachers cannot ignore examinations, it is difficult to subscribe to purely formative self-assessment (Carless, 2011; Xu, 2019), planning for the coexistence of summative and formative self-assessment is essential.

More importantly, teachers are advised to encourage students to share the responsibility for managing learning and assessment through engaging students in self-assessment activities. If students begin to believe that learning is not a competition but a pursuit of knowledge, they may feel empowered and confident to undertake self-assessment tasks realistically (Earl & Katz, 2006; González-Betancor et al., 2019). During the intervention, writing learning journals were used as a reflective tool for the case study students to understand their own learning process and deepen the quality of their self-assessment and critical thinking, as evidenced in previous studies (e.g., Boud, 2001; Wang, 2016). Understandings from journals can identify realistic approaches, as well as facilitating and constraining factors when students self-assess their EFL writing.

8.3 Limitations and Recommendations

Although this study collected data using a range of instruments to realise its aims, as with all research, it still has several limitations that need to be stated, and I believe that those limitations can help identify questions requiring future research.

8.3.1 Limitations of the research

Considering the current research findings, I need to point out two limitations. The first limitation arises from the instrument validation and data analysis approaches. Firstly, the decision to remove three items (6-8) from the *Self-efficacy for Self-assessment of EFL Writing* questionnaire based on factor analysis may lead to underrepresentation of the construct of writing, as those items are central to the organisation and ideas in the writing process. Therefore, the construct was only validated in a statistical sense. Secondly, findings of this study would have been much clearer if using mixed analysis of variance to present between-subjects factors (group) and within-subjects factors (time) with students background information (e.g., gender and years of studying English) considered as control variables.

Another limitation relates to sampling strategies. First, because the lecturer participants and the English-major student participants were obtained through convenience sampling from only one Chinese university, the findings of this research need to be interpreted with caution. Secondly, since only six students participated in the case study, the findings of this study may not be representative for generalisation of EFL students' experiences in self-assessment of writing to other populations. In addition, the EFL lecturer's experiences and cognitions about teaching self-assessment of writing are not explored, which meant the experiences in the classroom were not captured fully.

Therefore, caution should be taken when applying the findings of this research to other educational settings, year levels, or majors in universities.

8.3.2 Recommendations for future research

Despite the limitations as acknowledged above, I recommend that this study be replicated in other EFL settings with a comprehensive research design. The following are specific areas for further research.

A productive area for future research is to compare the effects of using self-assessment and teacher-assessment on students' writing performance in EFL writing settings, particularly, how to ensure the implementation of high-quality self-assessment in writing programmes (Laveault & Allal, 2016). Many researchers have contended that sophisticated self-assessment takes time, and the intervention time for this study may have been inadequate to document more systematic changes in the self-assessment process. Longitudinal studies over a much longer time span, are called for to provide richer insight into the effectiveness of the self-assessment intervention (Lee, 2011a; Meihami & Razmjoo, 2016). More in-depth qualitative studies adopting different methodologies to enrich the understanding of students' lived experiences in self-assessment of writing, are anticipated. For instance, more attention could be given to interpreting the mental and emotional reactions that students experience during self-assessment; further studies on the factors found in this study (students' English proficiency levels, previous teachers' feedback approach and cultural factors) that may affect EFL students' effective self-assessment of writing are also needed. Likewise, further studies using think-aloud protocols may lead to timely information on how individuals' internal cognition and affective mechanisms function in students'

engagement in self-assessment and how students' course of development in self-assessment may differ due to their personal differences.

Another interesting avenue for future research may be to expand the sampling methods and participant pool, such as students of different majors and ethnicities, to ensure the representation of the varied population in self-assessment research. For instance, the question can thus be raised: Would similar results be obtained if the self-assessment treatment were replicated with students from different year levels in an EFL context?

An investigation in China, into whether the consistent use of self-assessment within the EFL writing classrooms could strengthen students' confidence in self-assessment of writing, improve writing performance and influence their views and experiences on self-assessment of writing, also merits research attention.

It is also worth bearing in mind that it is simply not enough for students to develop their self-assessing confidence and capability only in the writing class. A more comprehensive understanding of self-assessment in second language education requires research into the development of self-assessment competence beyond the current programme of study.

Given that a few studies have already revealed misalignment between student and teacher views on self-assessment in different year levels (Panadero, Brown & Courtney, 2014; Wong, 2016), further studies are needed to investigate the nature and benefits of self-assessment of writing with both students and teachers. Research is required to understand how teachers and students could work together better to develop effective forms of self-assessment to maximise the benefits gained from the self-assessment of writing practices.

8.4 Conclusions

With the overarching aim in this research being to explore the effect of using self-assessment on EFL students' self-efficacy beliefs and writing performance, I advanced four main research questions, which are important constituents of students' lived experiences in self-assessment. Based on the findings and the discussion presented in the thesis, the following conclusions can be drawn.

Firstly, with the empirical evidence provided in the previous chapters, it is possible to conclude that self-assessment can be oriented to foster and enhance students' self-efficacy for self-assessment of writing and writing capacity and performance by assisting students' active engagement in learning. In this study, however, students' self-efficacy for self-assessment of writing did not emerge as a reliable predictor of their writing performance, perhaps due to the short intervention period.

Secondly, given that self-assessment was a novel experience for both the lecturer and Chinese students in this study, the inclusion of self-assessment in the writing classroom may take longer to have an effect on certain aspects, for instance, in improving students' rating accuracy and advancing students' writing complexity. Through this study, it has been demonstrated that the effectiveness of self-assessment practices might be contingent on individual differences; in the case studies, to some extent, students from diverse English proficiency levels reacted differently to self-assessment of writing during the intervention. Nevertheless, it is imperative to acknowledge and to conclude that, through regular self-assessment of writing practices, Chinese students found a sense of ownership of their writing by the end of the intervention. The insights gained from this study could assist in research and teaching writing in similar EFL contexts.

This study presents only the tip of a very promising iceberg, namely, students' self-assessment in the EFL writing context. I hope, therefore, the findings will encourage practitioners to implement, and monitor, self-assessment in their classroom despite the challenges and difficulties that may arise during their teaching practices.

APPENDICES

Appendix A: Examples of Participant Information Sheet and Consent Form

PARTICIPANT INFORMATION SHEET

Project Title: **The Effects of Using Self-assessment on English-as-a-Foreign-Language (EFL) Students' Self-efficacy and Writing Improvement**

Name of Researcher: Xiaoyu Zhang

Name of Supervisors: Professor Lawrence Jun Zhang, Professor Judy Parr, and Dr. Christine Biebricher.

Researcher Introduction

My name is Xiaoyu Zhang, and I am a Ph.D. candidate at the Faculty of Education and Social Work, The University of Auckland, New Zealand. I am conducting research on the effects of training English-as-a-foreign-language (EFL) students in self-assessment on their self-efficacy and writing improvement.

Project Description and Invitation

This project aims to develop EFL students' writing self-assessment ability through a self-assessment-based intervention. The effects of such teaching intervention on students' writing self-assessment confidence and writing performance will be explored. This study firstly aims to explore the current situation of Chinese EFL students' confidence in writing self-assessment, as well as their current writing self-assessment abilities. Then, it intends to carry out self-assessment skills-based writing intervention in the regular writing classrooms.

In total, this research plans to recruit 120 student participants. You are genuinely invited to join my Ph.D. research. I have contacted your faculty and obtained the Dean's permission to ask for your participation.

Faculty Involvement

Initially, I am seeking your consent to approve the research conduction in your English department. Also, I am seeking your help to approach EFL lecturers on my behalf and describe the research to EFL lecturers in a faculty meeting and asking their assistance to circulate the research information to students (distribute the Participant Information Sheets (PISs) and the Consent Forms (CFs). I will return to your site to collect these forms later if anyone is interested in the research. Secondly, I would like to request the usage of a spare classroom for interviews. All the potential participants will be informed of the complete research procedures in detail. I would also like to obtain your assurance that either participation or non-participation will not affect lecturers' or students' study,

academic performance or grade, future graduation or employment, and career development in any way. Additionally, I would like to give my assurance that those who do not participate will not be disadvantaged or underprivileged in any way.

Student Involvement

Participants from four intact classes will be randomly assigned to an intervention group (two classes) and a comparison group (two classes).

During the research period (09/2017-01/2018), both groups will be invited to complete the following in the first week (pre-test) and the last week (post-test) of the semester in your class time:

- Two English questionnaires focusing on students' confidence in writing self-assessment and practice. The questionnaire takes 20–30 minutes to complete, and it will be delivered to you in person at the beginning and again at the end of the research. It is required that you write a unique identification name, which can only be identified by yourself.
- Two 'cause and effect' writing samples based on your learning or your life (45mins) in class. For each of the writing sample, you are required to complete a given-topic essay (two similar topics, e.g.: What are the reasons for people shopping online and what effects it has on people's lives and the society? b. What are the reasons for people shopping in physical stores and what effects it has on people's lives and the society?) of at least 200 words within 45mins with a pen on the paper. Writing topics in the same difficulty and similarity level will be selected from the Chinese National Test for English Majors, Band 4 (TEM 4) past papers to ensure topic fairness in that familiarity should be similar for all participants. The outcome of the two writing samples is only for research purposes. They will not be related to your course performance.
- Self-assessing your writing samples with the help of the provided self-assessment writing rubric (20mins). You are required to self-assess your text in various aspects according to the writing rubric given to you, and comment on your challenges and strengths in writing, and to generate individual section score and an average overall score.

The intervention group participation (09/2017–01/2018) will include completing a 16-week writing intervention in their regular writing class time. All the students from the intervention group will receive:

- Writing instruction, in which participants from the intervention group will receive writing self-assessment skills-based instruction (one session per week of 1.5 hours) from their writing lecturer, as an intervention for around 16 weeks.

During the writing intervention, six students from the intervention group will be required to:

- Keep learning journals with pen and paper in the 4th, 10th and 16th week of the intervention (approximately 30 minutes to complete each) about changes concerning your use of self-assessment skills in the writing class, and your self-efficacy in writing self-assessment. The researcher will provide students with

prompts and framework to complete the journal, and then collect the journals during the research period with students' approval.

The non-participating students from the intervention group will join in the writing classes of the comparison group, where the same regular writing courses required by the university curriculum are taught, except the training of writing self-assessment. For the sake of equity, the researcher will provide the comparison group's lecturer with the self-assessment skills-based teaching package to implement it in classes after the intervention group completes the intervention programme. For anyone who does not take part in the intervention, it is assured that he/she will not be disadvantaged by this and their course grade will not be influenced.

Additionally, during the writing intervention, (09/2017–01/2018), six participants from the intervention group will be invited to attend a follow-up case study. You will participate in:

- Monthly semi-structured interviews, in which each may last approximately 30 minutes and will be audio-recorded. You are free to answer or to not answer any questions during the interview, and to request the recorder to be switched off at any time. The audio will be transcribed by the researcher. You will have two weeks to check if you would like to make any adjustments (to add/delete any statements) when you receive a copy of the completed transcription.

Lecturer Involvement

Before the conduction of this project (09/2017–01/2018), lecturers will be informed about the research project and be asked to spread research information: the Participant Information Sheets (PISs) and the Consent Forms (CFs) to their students on behalf of the researcher.

If any lecturers are in charge of a writing course during the research period (09/2017–01/2018), they and their respective two classes might be invited to participate in the self-assessment-based writing intervention. Two voluntary lecturers and their respective two classes will be randomly allocated to an intervention group (two intact classes) and a comparison group (two intact classes).

During the intervention, the lecturer of the intervention group will teach his/her classes on how to conduct writing self-assessment in class for around 16 weeks. Also, the lecturer of the intervention group will be invited to join a four-day training workshop with the researcher (two hours per day) to upgrade her knowledge base of using self-assessment in EFL writing. Moreover, to deal with the difficulties the lecturer encountered in implementing writing self-assessment in the previous weeks. The lecturer of the comparison group will follow his/her regular teaching procedure, and not applying self-assessment skills teaching in the writing classes. To ensure the implementation fidelity, the intervention group and comparison group teacher are required not to share their teaching practices, lesson plans, and other related materials during the time when the study was implemented in case they were influenced by each other. For the sake of equity, the researcher will provide the comparison group's lecturer with the self-assessment-based teaching package to implement it in classes after the intervention group completes the intervention programme.

In addition, during the intervention period, both the intervention group and the comparison group classes will be observed four times (1.5 hours each). The researcher will observe their teaching practices according to the classroom observation checklist, aiming to obtain information about the self-assessment skills-instructional process, classroom activities and students' engagement in the intervention group. The researcher will shortly interview the lecturer afterwards about how he/she feels during the writing class. Also, to ensure the comparison group is not using self-assessment in the writing classes. The researcher will share the observation checklist with the lecturers if they need it.

The researcher will send a summary of the research findings to the lecturers after the project if you would like one. If you are interested and would like to join this research, please give your consent to be contacted by the researcher on the attached CF.

Your participation or non-participation in the intervention or journal writing will not influence your relationship with your English lecturer, faculty and the university.

Anonymity and Confidentiality

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. Information about the research site: the university and the faculty will be disguised. Each participant will be given a digital code ranging from 001 to 092 by the researcher as his/her identification. The participants will use this code throughout the research to mark their writing samples, questionnaires, journals, recordings and so on. All participants' information will be kept separately from the data and will only be known to the researcher. If the information participants provide is going to be reported/published, pseudonyms or the unique identification name will be applied to protect their identity. No recognizable information and collected data from the research will be released to a third party.

Participants' Rights to Withdraw

Participants have the right not to answer any specific question at any stage, and they are entitled to withdraw themselves or any data provided by them at any time before 1st January 2018 without giving a reason. The Dean has provided assurances that participation, non-participation or withdrawal will not affect participants' normal courses, grades or relationships with the lecturers and faculty.

Data Storage, Retention, Destruction and Future Use

The collected hard copy data will be securely stored in a locked cabinet at The University of Auckland, and the electronic data will be stored on the researcher's password-protected computer in the University of Auckland. After six years, all hard copy data will be destroyed, and the digital information will be deleted. The data collected from the research will be used for the researcher's Ph.D. thesis at the University of Auckland, and may be also used for academic publications, and conference presentations. If you would like a copy of the research findings in the end, please indicate this on the consent form, and I will send a summary to you.

Thank you very much for taking the time to read this information sheet. If you have any inquiries or questions, please do not hesitate to contact me or my supervisors.

Contact details

Researcher	Main supervisor	Co-supervisor	Advisor
Xiaoyu Zhang School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland xiaoyu.zhang@auckland.ac.nz	Professor Lawrence Zhang Associate Dean International Strategic Engagement, School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland Gate 3, 74 Epsom Ave, Auckland. lj.zhang@auckland.ac.nz Ph: +64 9 6238899 ext: 48750	Professor Judy Parr School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland jm.parr@auckland.ac.nz Tel: 64 9 623 8899 ext 88998	Dr. Christine Biebricher School of Curriculum and Pedagogy, Faculty of Education and Social Work, The University of Auckland, Gate 3, 74 Epsom Ave, Auckland c.biebricher@auckland.ac.nz Tel: 6238899 ext. 48229

You may also contact the head of the School of Curriculum and Pedagogy, Associate Professor Helen Hedges at h.hedges@auckland.ac.nz or +64 9 373 7999 ext 48606.

Thank you very much for your time.

Yours sincerely,

Xiaoyu Zhang

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.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON…7th August 2017…………..For (3) YEARS, REFERENCE NUMBER………019574……….

CONSENT FORM

(THIS FORM WILL BE HELD FOR A PERIOD OF SIX YEARS)

Project Title: **The Effects of Using Self-assessment on English-as-a-Foreign-Language (EFL) Students' Self-efficacy and Writing Improvement**

Researcher: Xiaoyu Zhang

I have read the Participant Information Sheet and understood the nature of the research and why I have been invited to participate. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree to provide the research site for the research to be conducted in my department.
- I understand lecturers' and students' participation are voluntary.
- I agree to circulate the research information on the researcher's behalf to EFL writing lecturers and undergraduate students.
- I agree that the researcher join a faculty meeting to explain the research.
- I agree to allow EFL writing lecturers and undergraduate students in their corresponding classes to join this research.
- I understand that there is no grade evaluation during the writing intervention.
- I assure that participants' participation, non-participation or withdrawal will not affect their career, employment, grade, and relationships with the faculty and the university.
- I understand that student participants will be required to write a unique identification code, which can only be recognised by themselves on all items submitted during the whole research.
- I understand my participation is voluntary and I will be given a digital code ranging from 001 to 092 by the researcher as my identification throughout the research.
- I agree to take part in the 16 weeks writing intervention (once a week with 1.5 hours per session), and I understand that there is no grade evaluation on the writing intervention.
- I understand that I will be randomly assigned to either the intervention group, in which I will receive a 16-week writing instruction, or the comparison group, in which I will have the same period of the regular writing courses, and there is no bias in the process.
- I understand the intervention course will be available to the comparison group after the research project and if I miss the intervention course, I will not be disadvantaged by this.
- If I am part of the intervention group, I agree to keep journals during the intervention (4th, 10th and 16th week) and each of them may take 30 minutes to complete.
- I allow the researcher to access my writing journals.

- I understand I may be asked to join the follow-up case study.
- I understand my participation is voluntary and I agree to take part in the 40-minute face-to-face one-to-one semi-structured interviews before and after the intervention. I understand that the interviews will be conducted in my spare time during the term and will be audio-recorded.
- I understand that I have the right not to answer any specific questions in the interview, and I am entitled to have the recorder turned off at any stage without giving reasons.
- I understand that I will be invited to check the transcription and have the right to make any amendments and/or delete any statements in two weeks after I receive the transcript.
- I understand the interview is part of the follow-up case study of this research.
- I understand that if I miss any of the interviews, and I will not be disadvantaged by this.
- I understand that there is no grade evaluation in the interviews.
- If I were assigned to teach the intervention group, I will not share my teaching practices, teaching plans, and other related materials with the comparison group lecturer so that we would not be influenced by each other.
- I understand that I have the right to withdraw myself or any data provided by myself at any time before 1st January 2018 without explanation.
- I understand the Dean has assured that my participation, non-participation or withdrawal from this research will not affect my grade or relationship with my lecturers or faculty.
- 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 my anonymity may not be assured during the research, but the researcher will keep my identity confidential during the research process and that no identifying information will be included in any research outputs.
- I understand that the data collected from the research will be used for the researcher's Ph.D. 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 understand that information regarding the university and the faculty will be disguised, and 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 _____.

Name: …………………………… Date………………………...

Signature: …………………….…………………………………

Appendix B: Self-efficacy of Self-assessment in EFL Writing Questionnaire

(SESAEFLWQ-Finalised with 20 items)

You are invited to answer the following questions about yourself and your self-efficacy (the confidence of your ability) in writing self-assessment. This survey has two parts, and it is a part of a Ph.D. project at The University of Auckland. I would appreciate your honest answers. Thank you very much for your time and help.

Part One

In this part, please provide the following information by ticking (√) in the box or writing your response in the space so that we can interpret your answer better.

Code: _____ Gender: Male ☐ Female ☐ Gender diverse ☐
Age_____

University: _____ Major: _____

Length of English learning: _____ year(s) since ____ (age)

English level reflected by College entrance exam: _____ (score).

Overseas experience:

Have you ever been to English-speaking countries (e.g., studying, travelling, others)?

Yes ☐ No ☐ If Yes, how long? _____

For what? Travel ☐ Study ☐ Other (Please specify) _____

When do you self-assess your English writing?

In the writing course ☐ In out-of-class time ☐ Never ☐ Other (Please specify) _____

Part Two

Please answer the following questions concerning your confidence in your ability in English writing self-assessment. Indicate how much you agree or disagree with the following statements by circling a number from 1 to 5. There are no right or wrong answers; I am interested in your real situation and opinions. Please answer all questions by giving your best estimate.

Not at all sure I can do	Slightly sure I can do	Moderately sure I can	Very sure I can do	Completely sure I can do
1	2	3	4	5

For example:

I can tell what a good piece of argumentative writing likes.　　　　1 2 3 4 ⑤

1. I can identify incorrect words' spellings in my writing or other writing	1 2 3 4 5
2. I can assess if I use nouns, verbs, adjectives, etc. in the right way in my	1 2 3 4 5
3. I can assess if I use plurals, verb tenses, prefixes, and suffixes in the right	1 2 3 4 5
4. I can assess whether I write a simple sentence with the right grammar or not.	1 2 3 4 5
5. I can assess whether I write compound and complex sentences with the right	1 2 3 4 5
6. I can point out my strengths and weaknesses in my writing.	1 2 3 4 5
7. I assess whether I have clear goals before writing or not.	1 2 3 4 5
8. I can assess whether I realise my goals to improve my writing using different	1 2 3 4 5
9. I can revise my writing with the feedback from myself.	1 2 3 4 5
10. I can revise my writing when the teacher tells me what to do with my self-	1 2 3 4 5
11. I can self-assess my writing nicely with a detailed writing standard.	1 2 3 4 5
12. I can self-assess my writing according to the descriptions in writing standard of different	1 2 3 4 5
13. I can evaluate whether I achieve my goals in writing.	1 2 3 4 5
14. I can self-assess my writing process and point out the steps that I am not	1 2 3 4 5
15. I can estimate my writing performance correctly in class practices,	1 2 3 4 5
16. I can self-assess my writing with different methods if I were taught about them in the writing course.	1 2 3 4 5
17. I can self-assess my writing well if I learn how to do writing self-assessment	1 2 3 4 5
18. I can self-assess my writing well if I practice writing self-assessment	1 2 3 4 5
19. I can self-assess my writing better if I develop the writing standard on my	1 2 3 4 5
20. I can master writing self-assessment skills and self-assess my writing efficiently if my teacher encourages and supports me all the time.	1 2 3 4 5

Appendix C: Writing Topics in Part One

For the Pre-test (45 minutes)

Prompts: Shopping online is a new trend throughout the world today, and the reasons why people choose to shop online is complex. Such a trend may transform people's lives, or even how the society operates.

Write an essay with at least 200 words on the following topic: *What are the reasons for people shopping online and what effects does it have on people's lives and the society?*

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

For the Post-test (45 minutes)

Prompts: Even though shopping online is popular around the world, a large number of people still prefer shopping in physical stores due to different reasons, and how shopping in physical stores could shape people's lives and the whole society is not clear.

Write an essay with at least 200 words on the following topic: *What are the reasons for people shopping in physical stores and what effects does it have on people's lives and the society?*

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

Appendix D: Self-assessment Writing Rubric

There are four sections in this writing rubric, namely, task achievement, coherence and cohesion, language resources, and mechanics. Within each section are descriptions of characteristics representing writing at different levels, from 1-5, where 5 represents the highest level of quality. I would like to request you to self-assess your writing according to the rubric, and generate a score for yourself in each section and an average overall score. Finally, please write down your comments on the strengths and weaknesses of your writing. You have 20 minutes to complete this task. The meanings or definitions of the key dimensions in the writing rubric are provided as follows. I suggest that you read through the rubric to become familiar with it before starting. Thank you very much for your generous help.

- *Task achievement*: That means the extent to which you respond to all parts of the topic questions, and how you develop relevant ideas to support them with explanations, examples or experience. It is important that your opinion is clear and relevant to the topic.
- *Coherence and Cohesion:* Coherence means that you present your ideas in a logical way, and cohesion refers to the degree you connect sentences (or different parts of one sentence) and paragraphs with linking words (e.g., since/in addition/ because/first/ although/however/moreover etc.) so that the reader could follow your ideas easily. Both of them build up the organisation of your writing.
- *Language Resources:* This refers to your ability to use a variety of vocabulary (e.g., different word forms: noun/verb, synonyms, phrases, formal expressions), grammar (e.g., different tenses: I had gone/I will be going, and relative clauses: who/which/that etc.) and sentence structures (e.g., simple sentences—one main clause/one verb; complex sentences-- at least one independent clause plus at least one dependent clause linked by because, although etc.; compound sentences--two or more independent clauses joined by and, but, or) in a balanced, flexible and accurate way in your writing.
- *Mechanics:* This is your ability to apply correct spelling, punctuation (e.g., comma, period, and question mark), capitalization, abbreviations (e.g., D.I.Y, CEO), paragraphing (divide your writing into logical parts), and grammar rules (e.g., the use of the/a/an, uncountable nouns, plural forms, use of phrases match, avoiding run-on sentences: in which two or more independent clauses (i.e. complete sentences) are joined without an appropriate conjunction or mark of punctuation.) in your writing.

I believe my strengths in this writing are:

_____

_____

I think my weaknesses in this writing are:

_____

_____

My scores are □ task achievement, □ coherence and cohesion, □ language resources, and □ mechanics. My general score is □

Level	Task Achievement	Coherence and Cohesion	Language Resources	Mechanics
5	☐ Proficiently paraphrase and address all the task questions. ☐ Clearly express your opinions through the task and support them with at least four relevant, detailed examples, and personal experiences. Meet or moderately exceed the word limit.	☐ Ideas, sentences, and paragraphs are clearly and logically presented and linked by a variety of linking words accurately. ☐ With topic and closing sentences, the whole writing has a nice flow and pleasant to read.	☐ Use a wide range of topic-related and sophisticated vocabulary naturally to answer task questions; may have rare mistakes in word choice, and forms. ☐ Use a wide variety of sentence structures, and all are used in a flexible and correct way.	☐ No errors in spelling, punctuation, capitalization, grammar rules. ☐ Divide the writing into reasonable paragraphs, and each paragraph expresses a clear and logical idea and indicates a change of focus at the beginning. Clear and neat handwriting.
4	☐ Appropriately paraphrase and address about 85% of the task questions. ☐ Clearly express your opinions through the task and support them with three generally relevant, detailed examples, and personal experiences, but they may lack focus sometime. Meet or slightly exceed the word limit.	☐ Though 1-2 expressions are unclear, ideas, sentences, and paragraphs are are clearly and logically presented in general. Some linking words are underused or overused. ☐ With topic and closing sentences, the whole writing is easy to follow.	☐ Use a wide range of topic-related and sophisticated vocabulary to answer task questions; may have 1-2 mistakes in word choice, and word forms. ☐ Use a wide variety of sentence structures, and about 85% of them are error-free sentences.	☐ 1-3 errors in spelling, punctuation, capitalization, grammar rules. ☐ Divide the writing into reasonable paragraphs, and each paragraph expresses a central idea and indicates a change of focus at the beginning. Clear and neat handwriting.
3	☐ Clearly address and paraphrase around half of the task questions. ☐ Express a relevant but unclear personal opinion through the task and support it with at least one relevant, detailed example and personal experience. Slightly or moderately below the word limit.	☐ Parts of ideas, sentences and paragraphs are repetitive and illogical. Different linking words are used, but there are 3-4 mistakes within/between sentences that may block reading but do not affect communication. ☐ The writing has topic sentences for each part, but no conclusion.	☐ Use basic topic-related vocabulary to answer task questions, but 3-5 errors in word forms or grammar of uncommon words. ☐ Use a mix of simple and complex sentence structures; simple sentence structures are used correctly, while having 2-4 mistakes in complex sentences.	☐ 4-7 noticeable errors in spelling, punctuation, capitalization, grammar rules. ☐ Divide the writing into reasonable paragraphs, but each paragraph expresses more than one idea. Good handwriting without affecting reading.
2	☐ Unable to address any part of the task questions, and almost copy all the words of the task questions. ☐ Express an unclear personal opinion through the task but the supporting examples, and personal experiences are irrelevant and not well-explained. Considerably below the word limit.	☐ It is hard to follow the logic in ideas, sentences, and paragraphs. Only basic linking words (e.g., and/or/but/first/however) are used, and they may be used incorrectly and repetitively. ☐ The writing does not have a clear topic sentence and a conclusion.	☐ Use only basic vocabulary and some of them are not related to the task. Over 6 errors in word forms and the meaning of ideas may be changed by the errors. ☐ Use only simple sentences without clauses. Many incomplete and run-on sentences.	☐ Over 8 errors in spelling, punctuation, capitalization, grammar rules that the writing is hard to read and understand. ☐ There are less than 3 paragraphs or more than 5 paragraphs. Ideas for different questions are mixed so the meaning is confusing for the reader. Poor handwriting, difficult to read.
1	☐ The answer is mostly unrelated to the task, no personal opinion is expressed and it seems part of the writing is memorized response. Extremely below the word limit.	☐ No logic in the presented ideas, sentences, and paragraphs; wrong use of all the linking words; no beginning or ending.	☐ Use extremely limited and repetitive vocabulary, and most of the word forms are used wrongly; only run-on sentences or phrases which are hard to read.	☐ With very poor handwriting, the writing is full of errors and no paragraph formatting. Almost impossible to read.

Appendix E: Classroom Observation Checklist

Lecturer: _____Course: _____Year: _____Date: _____

Students No.: _____

Teaching content	
Course material	
Teacher support/facilitation in class	
The assignment given and evaluation	
Classroom activities	

Other course-related issues

Instructional Process (Teaching sequence and moves)

Self-assessment teaching strategies	Note:
• Explicit, step-by-step explanation of using self-assessment in writing. Yes ☐ No ☐ • Explaining criteria for students to use in self-assessment of writing. Yes ☐ No ☐ • Exemplars use to demonstrate self-assessment of writing. Yes ☐ No ☐ • Guidance on self-assessing writing components and the whole writing. Yes ☐ No ☐ • Students' participation in self-assessment criteria development. Yes ☐ No ☐ • In-class writing self-assessment exercises without grading. Yes ☐ No ☐ • Student-teacher dialogue during self-assessment exercises. Yes ☐ No ☐ • Time for students' to revise writing after self-assessment. Yes ☐ No ☐ • Encouragement, support and timely feedback about students' self-assessment. Yes ☐ No ☐	

Appendix F: Sample Lesson Plan for Self-assessment-based Writing

Intervention

Lesson 1

Teaching objectives:

To implement effective self-assessment in classrooms, initially, this lesson aims to retrieve and activate students' previous knowledge of self-assessment; to raise students' awareness about the importance of self-assessment; to demonstrate the approaches and tools can be used for self-assessment of writing with detailed explanation (Nielsen, 2019; Rolheiser, 1998; Ross et al., 1999). In addition, informed by Andrade and Brown's (2016, p. 327) study, which highlighted three means to facilitate students to perform self-assessment, in the current study, the teacher for the treatment group provided students "how-to instruction, authentic engagement, and practice with feedback in an environment of trust".

Teaching procedures:

1. Providing students a checklist of questions to assess their own previous knowledge and practices about self-assessment of writing.

1) What does self-assessment mean?
2) What does self-assessment mean to me?
3) Why do I need self-assessment (for my writing)?
4) Think back, have I self-assessed myself/my learning before?
5) If I have self-assessed before, how did I self-assess myself/my learning?
6) If I have not self-assessed before, how would I self-assess myself/my learning if given the opportunity?
7) What can I achieve from self-assessment?
8) What are the sources I can use to help my self-assessment of writing?

2. Asking students to share their reflections on their writing and the above questions in the form of oral presentations so that the teacher could gain valuable insight into students' understanding and readiness for moving into self-assessment-based writing intervention.
3. Developing students' background knowledge of expository writing, and introducing how self-assessment would fit in the entire writing process, namely, pre-writing, during writing and post-writing.
4. Guiding students to set personal achievable goals for this class and to keep a record of the goals themselves.
5. Introducing self-assessment writing rubric, and examining each level and dimension of the rubric with students together to clear students' uncertainties when using the rubric.
6. Guiding students to self-assess paragraph exemplars using the rubric provided and discussing possible approaches to improve those exemplars with students.

7. Providing students with sufficient time to reflect on the reasons for whether they have or have not achieved the goals they set at the beginning of the class before the class finished.
8. Encouraging students to self-assess their assignment essay using the provided rubric when they finished it and commenting on their own essay.

Lesson 2

Teaching objectives:

To ensure students that they are in a secure and friendly environment in which their self-assessment is for personal growth only; to enhance students' confidence in self-assessing their writing products by providing in-time feedback and encouragement regarding their self-assessment; to modal self-assessment of longer passages using the writing rubric (Earl & Katz, 2006; Harris & Brown, 2018).

Teaching procedures:

1. Self-reflection: students answer questions such as what have I learned about expository writing last class? What are my goals for this writing class?
2. Analyse and evaluation of exemplars: students assessed the writing exemplars according to the writing rubric provided and presented their opinions of the exemplars, the rubric, and how the rubric and exemplars could be possibly improved
3. Self-checklist: students were encouraged to develop their own checklist of questions for the whole writing process when they received the writing topic. For instance, Have I brainstormed ideas for this essay? Have I prepared an outline for this essay? Have I written clear topic sentences for this essay?
4. Self-reflection: students read their completed essays individually and silently; then, students voluntarily described their writing process and identified their writing strengths and weaknesses to the teacher or their peers.
5. Interpretation of self-assessment results: The teacher commented on students' self-assessment in private in the form of one-to-one conferencing and inspired students to consider their writing critically with questions such as does the essay's opening fully address the topic? How can you make your argument more convincing?
6. Self-review: students were encouraged to review their goals for this class and adjust the goals if necessary.

Appendix G: Students' Interview Schemes for the Pre-test and the Post-test

1. Have you heard of self-assessment of writing?

2. What do you think self-assessment of writing means?

3. In any of your university courses, do you revise or reflect on your writing when you finish it? If yes, how do you go about this? Which aspects do you focus on?

4. How do you self-assess your English writing in your writing course? What do you focus on? How often (every piece of writing)?

5. Do you self-assess your English writing in out-of-class time? If yes, how and how often?

6. What kind of qualities do you think one needs to self-assess his/her own English writing?

7. Why do you think some students can self-assess their compositions but others not?

8. Do you think self-assess one's own English writing would help an individual to make progress in writing? Why do you say that? In what way might self-assessment help?

9. Do you like the way that your lecturer is teaching you about writing self-assessment in class? What makes you say this and come to this conclusion?

10. What do you think of the role of writing rubrics when you self-assess your writing? Could you explain why they are helpful/ not helpful/ how you use them specifically?

11. What do you think of the usefulness of the writing course such as you are taking? Do you have any comments on improving the writing instruction?

12. What have you learned or achieved since learning to self-assess your writing in the class?

13. What do you think are the difficulties (e.g. the hardest parts) when practicing writing self-assessment so far?

14. Which part do you like best during the self-assessment of writing?

Appendix H: Learning Journal Framework

Time 1 (4th week)

Last month, what have you learned about writing self-assessment?

_____
_____
_____

In the last month, what has changed in your understanding of writing self-assessment?

_____
_____
_____

When I self-assessed my writing last month,

I feel_____,
because_____
_____
_____

Currently, what I like best about writing self-assessment is

_____
_____
_____

One thing I would like to improve about my writing in the next month writing class is

_____
_____
_____

Time 2 (10th week)

Last month, what new things have you learned about writing self-assessment?

_____
_____
_____

When I self-assess my writing last month,

I feel_____,
because_____
_____

_____

Currently, what I like best about writing self-assessment
is_____
_____

One thing I would like to improve about my writing in the next month writing class
is_____
_____

Time 3 (the end of the fourth month)

Currently, what I like best about writing self-assessment
is_____
_____

After a semester time, what has changed in your understanding about writing self-assessment?

_____
_____
_____

Looking back, the most difficult part of writing self-assessment for me was
_____
_____

After learning and practicing writing self-assessment about a semester, I believe I was most successful at
_____
_____
because_____
_____

Appendix I: Descriptive Statistics of the Finalised Self-efficacy for Self-assessment of EFL Writing Questionnaire (20 Items)

Items	M	SD	Skewness	Kurtosis
Item 1-I can identify incorrect words' spellings in my writing or other writing samples.	3.35	.94	-.17	-.24
Item 2-I can assess if I use nouns, verbs, adjectives, etc. in the right way in my writing.	3.59	.94	-.42	-.01
Item 3-I can assess if I use plurals, verb tenses, prefixes, and suffixes in the right way in my writing.	3.34	.94	-.14	-.34
Item 4-I can assess whether I write a simple sentence with the right grammar or not.	4.18	.93	-.97	.36
Item 5-I can assess whether I write compound and complex sentences with the right grammar or not.	3.22	.93	-.19	-.18
Item 9-I can point out my strengths and weaknesses in my writing.	3.30	.95	-.17	-.39
Item 10-I can assess whether I have clear goals before writing or not.	3.34	.90	-.08	-.33
Item 11-I can assess whether I realise my goals to improve my writing using different methods or not.	2.99	.85	.23	-.02
Item 12-I can revise my writing with the feedback from myself.	3.04	.96	-.06	-.34
Item 13-I can revise my writing on my own when the teacher tells me what to do with my self-assessment.	3.68	.87	-.31	-.20
Item 14-I can self-assess my writing nicely with a detailed writing standard.	3.20	.86	-.06	-.22
Item 16-I can assess my writing according to the descriptions in writing standard of different writing styles.	2.97	.89	.07	-.29
Item 17-I can evaluate whether I achieve my goals in writing.	2.89	.86	.03	-.14
Item 18-I can self-assess my writing process and point out the steps that I am not satisfied with.	3.55	.87	-.31	-.27
Item 19-I can estimate my writing performance correctly in class practices, homework, and exams.	3.00	.84	.02	-.19
Item 20-I can self-assess my writing with different methods if I were taught about them in the writing course.	3.25	.87	.13	-.34
Item 21-I can self-assess my writing well if I learn how to do writing self-assessment in class.	3.39	.85	-.12	-.21
Item 22-I can self-assess my writing well if I practice writing self-assessment regularly in class.	3.59	.82	-.25	-.05
Item 23-I can self-assess my writing better if I develop the writing standard on my own.	3.36	.83	-.14	-.04
Item 24- I can master writing self-assessment skills and self-assess my writing efficiently if my teacher encourages and supports me all the time.	3.61	.90	-.23	-.29

REFERENCES

Abdel Latif, M. (2013). What do we mean by writing fluency and how can it be validly measured? *Applied Linguistics*, *34*(1), 99–105.

Abedi, J. (2010). Research and recommendations for formative assessment with English language learners. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 181–197). New York, NY: Routledge.

Alishah, A. R., & Dolmaci, M. (2013). The interface between self-efficacy concerning the self-assessment on students studying English as a foreign language. *Procedia Social and Behavioral Sciences*, *70*, 873–881.

Andrade, H., & Brookhart, S. M. (2016). The role of classroom assessment in supporting self-regulated learning. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (Vol. 4, pp. 293–309). Chem, Switzerland: Springer.

Andrade, H., & Du, Y. (2007). Student responses to criteria‐referenced self‐assessment. *Assessment & Evaluation in Higher Education*, *32*(2), 159–181.

Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research*, *97*(1), 21–30.

Andrade, H. L., & Brown, G. T. (2016). Student self-assessment in the classroom. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 319–334). New York, NY: Routledge.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, *48*(1), 12–19.

Andrade, H. L, Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice*, *27*(2), 3–13.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*(8), 1–13.

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice*, *17*(2), 199–214.

Andrade, H. L. (2010). Students as the definitive source of formative assessment:

Academic self-assessment and the self-regulation of learning. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). New York, NY: Routledge.

Andrade, H. L, Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, *102*(4), 287–302.

Assessment Reform Group. (2002). *Assessment for learning: 10 principles (leaflet printed by the ARG).* Retrieved from: https://www.lancsngfl.ac.uk/secondary/math/download/file/PDF/AfL%2010%20principles.pdf

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, *6*(1), 14–29.

Bai, B., Chao, G. C. N., & Wang, C. (2019). The relationship between social support, self-efficacy, and English language learning achievement in Hong Kong. *TESOL Quarterly*, *53*(1), 208–221.

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

Bandura, A. (1986b). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359–373.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*(9), 1175–1184.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Charlotte, NC: Information Age.

Birjandi, P., & Hadidi Tamjid, N. (2012). The role of self, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, *37*(5), 513–533.

Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: A comparative study of their effect on writing performance and rating accuracy. *Iranian Journal of Applied Linguistics*, *13*(1), 23–45.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8–21.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills : Implications for teachers and researchers. *Language Learning*, *39*(3), 313–340.

Bloom, B. S., Hasting, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York, NY: McGraw-Hill.

Blue, G. M. (1988). Self-assessment—The limits of learner independence. In A. Brookes & P. Grundy (Eds.), *Individualization and autonomy in language learning: ELT documents 131* (pp. 100–118). London, England: Modern English Publications.

Boscolo, P., & Hidi, S. (2007). The multiple meaning of motivation to write. In S. Hidi & P. Boscolo (Eds.), *Writing and motivation* (pp. 1–14). Oxford, England: Elsevier.

Boud, D. (1989). The role of self-assessment in student grading. *Assessment & Evaluation in Higher Education*, *14*(1), 20–30.

Boud, D. (1995a). Assessment and learning: Contradictory or complementary. In P. Night (Ed.), *Assessment & learning in higher education* (pp. 35–48). London, England: Kogan Page.

Boud, D. (1995b). *Enhancing learning through self-assessment*. New York, NY: Routledge.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education*, (90), 9–17.

Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, *18*(5), 529–549.

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, *41*(3), 400–413.

Bourke, R. (2014). Self-assessment in professional programmes within tertiary institutions. *Teaching in Higher Education*, *19*(8), 908–918.

Bouziane, A., & Zyad, H. (2018). The impact of self and peer assessment on L2 writing: The case of moodle workshops. In A. Ahmed & H. Abouabdelkader (Eds.), *Assessing EFL writing in the 21st century Arab world: Revealing the*

unknown (pp. 111–135). Cham, Switzerland: Palgrave Macmillan.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77–101.

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic analysis. In P. Liamputtong (Ed.), *Handbook of research methods in health social sciences* (pp. 1–18). Singapore: Springer.

Broadfoot, P. M., Daugherty, R., Gardner, J., Gipps, C. V., Harlen, W., & James, M. (1999). *Assessment for learning: Beyond the black box*. Cambridge, UK: University of Cambridge Faculty of Education.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria: ASCD.

Brown, A. (2005). Self-assessment of writing in independent language learning programs: The value of annotated samples. *Assessing Writing*, *10*(3), 174–191.

Brown, G. T. L., & Harris, L. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, *3*, 22–30.

Brown, G. T. L., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education: Principles, Policy & Practice*, *22*(4), 444–457.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J.H. McMillan (Ed.), *SAGE Handbook of research on classroom assessment* (pp. 367–393). SAGE Publications.

Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). New York, NY: Pearson Education.

Bruning, R., Dempsey, M. S., Kauffman, D. F., Mckim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, *105*(1), 25–38.

Bruning, R. H., & Kauffman, D. F. (2015). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 160–173). New York, NY: Guilford Press.

Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, *60*(6), 626–648.

Butler, Y. G. (2018). Young learners' processes and rationales for responding to self-

assessment items: Case of generic can-do and five-point, likert-type formats. In M. J. Davis, J. M. Norris, M. E. Malone, T. H. McKay, & Y. A. Son (Eds.), *Useful assessment and evaluation in language education* (pp. 21–40). Washington, DC: Georgetown University Press.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, *27*(1), 5–31.

Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York, NY: Routledge.

Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing*, *44*(4), 51–62.

Cassidy, S. (2007). Assessing "inexperienced" students' ability to self-assess: Exploring links with learning style and academic personal control. *Assessment and Evaluation in Higher Education*, *32*(3), 313–330.

Chen, J., & Zhang, L. J. (2019). Assessing student-writers' self-efficacy beliefs about text revision in EFL writing. *Assessing Writing*, *40*(3), 27–41.

Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, *12*(2), 235–262.

Clark, I. (2012). Formative assessment : Assessment is for self-regulated learning. *Educational Psychology Review*, *24*(2), 205–249.

Cliffs, E., Hall, J., & Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248–287.

Coakes, S. J., Steed, L., & Price, J. (2008). *SPSS analysis without anguish: Version 15.0 for Windows*. Milton, Australia: Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Connelly, L. M. (2009). Mixed methods studies. *Medsurg Nursing*, *18*(1), 31–32.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, *10*(7), 1–10.

Cowie, N. (2009). Observation. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 165–182). New York, NY: Palgrave Macmillan.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative,*

and mixed methods approaches (5th ed.). Thouand Oaks, CA: SAGE.

Crooks, T. (2011). Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation*, *37*(1), 71–77.

Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. London, England: Routledge/Falmer.

Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 149–166.

DeVellis, R. F. (2012). *Scale development: Theory and applications.* Washington, DC: SAGE.

Ding, Y., & Zhao, T. (2019). Chinese university EFL teachers' and students' beliefs about EFL writing: Differences, influences, and pedagogical implications. *Chinese Journal of Applied Linguistics*, *42*(2), 163–181.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, *24*(3), 331–350.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge, England: Cambridge University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford, England: Oxford University Press.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York, NY: Routledge.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest, Supplement*, *5*(3), 69–106.

Earl, L., & Katz, S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Retrieved from: http://www.wncp.ca/media/40539/rethink.pdf

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximise student learning.* (2nd ed.). Thousand Oaks, CA: Hawker Brownlow Education.

Ellis, R., & Barkhuizen, G. P. (2005). *Analyzing learner language*. Oxford, England: Oxford University Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999).

Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299.

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A users' guide. *International Journal of Selection and Assessment, 1*(2), 84–94.

Field, A. P. (2013). *Discovering statistics using SPSS* (4th ed.). Thousand Oaks, CA: Sage.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2015). *How to design and evaluate research in education* (9th ed.). New York, NY: McGraw-Hill Education.

Gardner, D. (2000). Self-assessment for autonomous language learners. *Links & Letters*, *7*, 49–60.

Gardner, D., & Miller, L. (1999). *Establishing self-access: From theory to practice.* Cambridge, England: Cambridge University Print.

Gil, J. (2016). English language education policies in the People's Republic of China. In R. KirKpatrick (Ed.), *English language education policy in Asia* (Vol. 11, pp. 49–90). Cham, Switzerland: Springer.

Glasswell, K., Parr, J., & Aikman, M. (2001). Development of the asTTle writing assessment rubrics for scoring extended writing tasks. *University of Auckland, AsTTle Project*, *6*, 1–27.

González-Betancor, S. M., Bolívar-Cruz, A., & Verano-Tacoronte, D. (2019). Self-assessment accuracy in higher education: The influence of gender and performance of university students. *Active Learning in Higher Education*, *20*(2), 101–114.

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development*, *20*(1), 53–70.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, *36*, 101–111.

Harris, L. R., & Brown, G. T. L. (2018). *Using self-assessment to improve student learning*. New York, NY: Routledge.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, *51*(1), 12–20.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to*

achievement. London, UK: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Hawe, E., & Parr, J. (2014). Assessment for Learning in the writing classroom: An incomplete realisation. *The Curriculum Journal*, *25*(2), 210–237.

Henner-Stanchina, C., & Holec, H. (1985). Evaluation in an autonomous learning scheme. In P. Riley (Ed.), *Discourse and learning.* Chicago, IL: Longman Publishing Group.

Hobson, E. H. (1996). Encouraging self-assessment: Writing as active learning. *New Directions for Teaching and Learning*, *67*, 45–58.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, *17*, 63–84.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Huang, S. C. (2016). Understanding learners' self-assessment and self-feedback on their foreign language speaking performance. *Assessment & Evaluation in Higher Education*, *41*(6), 803–820.

Li, H. H., Zhang, L., & 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* (570575), 1–16.

IELTS (2015). *IELTS task 2 writing band descriptors (public version)*. Retrieved from: https://takeielts.britishcouncil.org/sites/default/files/IELTS_task_2_Writing_band _descriptors

Jacobs, H., Zingraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach.* Roweley: Newbury House.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144.

Kearney, S. (2013). Improving engagement: The use of "Authentic self-and peer-assessment for learning" to enhance the student learning experience. *Assessment*

& *Evaluation in Higher Education*, *38*(7), 875–891.

Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *American Journal of Psychology*, *103*(3), 327–342.

Kim, D. H., Wang, C., Ahn, H. S., & Bong, M. (2015). English language learners' self-efficacy profiles and relationship with self-regulated learning strategies. *Learning and Individual Differences*, *38*, 136–142.

King, N. (2004). Using templaktes in the thematic analysis of text. In C. Cassell & S. Gillian (Eds.), *Essential guide to qualitative methods in organizational research* (pp. 257–270). London, England: Sage.

Kirby, N. F., & Downs, C. T. (2007). Self-assessment and the disadvantaged student: Potential for encouraging self-regulated learning? *Assessment & Evaluation in Higher Education*, *32*(4), 475–494.

Klenowski, V. (1995). Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England. *Assessment in Education: Principles, Policy & Practice*, *2*(2), 145–163.

Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 263–268.

Kline, R. B. (2016). *Principles and practices of structural equation modeling*. (4th ed.). New York, NY: Guilford.

Koch, T. (2006). Establishing rigour in qualitative research: The decision trail. *Journal of Advanced Nursing*, *53*(1), 91–103.

Kun, A. I. (2016). A comparison of self versus tutor assessment among Hungarian undergraduate business students. *Assessment & Evaluation in Higher Education*, *41*(3), 350–367.

Lam, R. (2013). The relationship between assessment types and text revision. *ELT Journal*, *67*(4), 446–458.

Lam, R. (2016). Assessment as learning: Examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, *41*(11), 1900–1917.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 579–589.

Laveault, D., & Allal, L. (2016). Implementing assessment for learning: Theoretical

and practical issues. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 1–16). Cham, Switzerland: Springer International Publishing.

Leach, L. (2012). Optional self-assessment: Some tensions and dilemmas. *Assessment and Evaluation in Higher Education*, *37*(2), 137–147.

LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, *19*(4), 673–687.

Lee, I. (2011a). Bringing innovation to EFL writing through a focus on assessment for learning. *Innovation in Language Learning and Teaching*, *5*(1), 19–33.

Lee, I. (2011b). Formative assessment in EFL writing: An exploratory case study. *Changing English: Studies in Culture and Education*, *18*(1), 99–111.

Lee, I. (2014). Revisiting teacher feedback in EFL writing from sociocultural perspectives. *TESOL Quarterly*, *48*(1), 201–213. https://doi.org/10.1002/tesq.153

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Cham, Switzerland: Springer.

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing*, *22*(1), 34–50.

Li, H., & Liu, R. (2013). Correlation between College students' English writing anxiety and writing self-efficacy and its predication of students' writing performance. *Foreign Language Research*, *2*(138), 48–54.

Lincoln, Y., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA:Sage.

Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, *20*(1), 71–94.

Liu, H., & Brantmeier, C. (2019). "I know English": Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System*, *80*, 60–72.

Liu, J. (2002). Self-assessment of English writing skills by Chinese university Students. *Modern Foreign Languages (Quarterly)*, *25*(3), 241–249.

Liu, J., & Xu, Y. (2017). Assessment for learning in English language classrooms in China: Contexts, problems, and solutions. In H. Reinders, D. Nunan, & B. Zou (Eds.), *Innovation in language learning and teaching: The case of China* (pp. 17–37). London, England: Palgrave Macmillan.

Liu, Q. (2010). The national education examinations authority and its English tests. In

L. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 29-43). New York, NY: Taylor & Francis Group,.

Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. London, England: Palgrave Macmillan Education.

Longhurst, N., & Norton, L. S. (1997). Self-assessment in coursework essays. *Studies in Educational Evaluation*, *23*(4), 319–330.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, *45*(1), 36–62.

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, *96*(2), 190–208.

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, *34*(4), 493–511.

Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, *26*(1), 75–100.

Mazloomi, S., & Khabiri, M. (2018). The impact of self-assessment on language learners' writing skill. *Innovations in Education and Teaching International*, *55*(1), 91–100.

McCarthy, P., Meier, S., & Rinderer, R. (1985). Self-efficacy and writing : A different view of self-evaluation. *College Composition and Communication*, *36*(4), 465–471.

McMillan, J. H, & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons*, *87*(1), 40–49.

Meihami, H., & Razmjoo, S. A. (2016). An emic perspective toward challenges and solutions of self- and peer-assessment in writing courses. *Asian-Pacific Journal of Second and Foreign Language Education*, *1*(9), 1–20.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). San Francisco, CA: John Wiley & Sons.

Ministry of Education. (2018). *Higher education: National syllabus and standards for English majors in the People's Republic of China.* Beijing, China: Foreign Language Teaching and Research Press

Mok, M. M. C., Lung, C. L., Cheng, D. P. W., Cheung, R. H. P., & Ng, M. L. (2006). Self-assessment in higher education: Experience in using a metacognitive

approach in five case studies. *Assessment & Evaluation in Higher Education*, *31*(4), 415–433.

Mujis, D. (2011). *Doing quantitative research in education with SPSS*. London, UK: SAGE Publications.

Nawas, A. (2020). Grading anxiety with self and peer-assessment: A mixed-method study in an indonesian EFL context. *Issues in Educational Research*, *30*(1), 224–244.

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: The case of China* (pp. 63–87). London, England: Palgrave Macmillan.

Nicol, D. J., & Macfarlane-dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

Nielsen, K. (2012). Self-assessment methods in writing instruction: A conceptual framework, successful practices and essential strategies. *Journal of Research in Reading*, 1–16.

Nielsen, K. (2019). Peer and self-assessment practices for writing across the curriculum: Learner-differentiated effects on writing achievement. *Educational Review*, 1–22.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*, 1–13.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396–402.

Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment & Evaluation in Higher Education*, *22*(4), 357–368.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college level L2 writing. *Applied Linguistics*, *24*(4), 492–518.

Oscarson, A. D. (2009). *Self-assessment of writing in learning English as a foreign language*. Goteborg, Sweden: Geson Hylte Tryck.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, *6*(1), 1–13.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, *19*(2), 139–158.

Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Measurement and Evaluation in Counseling and Development*, *39*(4), 239–249.

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111–139). New York, NY: Erlbaum.

Pajares, F., & Johnson, M. J. (1994). Confidence and competence in writing: The role of self-efficacy, outcome expectancy, and apprehension. *Research in the Teaching of English*, *28*(3), 313–331.

Pajares, F., Miller, M. D., & Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *Journal of Educational Psychology*, *91*(1), 50–61.

Pajares, F., & Urdan, T. (2005). *Self-efficacy beliefs of adolescents*. Greenwich, UK: IAP.

Pajares, F., & Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *The Journal of Educational Research*, *90*(6), 353–360.

Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology*, *24*, 390–405.

Pajares, F., & Valiante, G. (2015). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S.Graham , & J. Fitzgerald (Eds.) *Handbook of writing research* (2nd ed.) (pp. 160–173). New York, NY: The Guilford Press.

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. (6th ed.). Sydney, Australia: Allen & Unwinb.

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practical connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology*, *11*(2), 551–576.

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and

formative assessment: A roadmap of where we are, how we got here, and where we are going. *Australian Educational Researcher*, *45*(1), 13–31.

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education*, *34*(3), 535–557.

Panadero, E., Brown, G. T. L., & Courtney, M. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy & Practice*, *21*(4), 365–383.

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known, unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98.

Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 311–324). Cham, Switzerland: Springer International Publishing.

Paris, S. G., & Paris, A. H. (2001). Classrooom applications of research on self-regulated learning. *Educational Psychologist*, *36*(2), 89–101.

Patri, M. (2002). The influence of peer feedback on self and peer-assessment of oral skills. *Language Testing*, *19*(2), 109–131.

Pausch, R., & Zaslow, J. (2008). *The last lecture*. London, England: Hodder & Stoughton.

Petrić, B., & Czárl, B. (2003). Validating a writing strategy questionnaire. *System*, *31*(2), 187–215.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International Universities Press.

Pinner, R. (2016). Trouble in paradise: Self-assessment and the Tao. *Language Teaching Research*, *20*(2), 181–195.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *American Journal of Psychology*, *41*(4), 219–225.

Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, *47*(1), 101–143.

Powell, K. C., & Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education*, *130*(2), 241–250.

Powers, D. E., Fowles, M. E., Farnum, M., & Ramsey, P. (1994). Will they think less of my handwritten essay if others word process theirs? Effects on essay scores of intermingling handwritten and word-processed essays. *Journal of Educational Measurement*, *31*(3), 220–233.

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, *34*(1), 1–20.

Rallis, S. F., & Rossman, G. B. (2009). Ethics and trustworthiness. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 263–287). New York, NY: Palgrave Macmillan.

Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. New York, NY: Routledge.

Read, J. (2010). Comprehensive measures of vocabulary. In *Assessing vocabulary* (pp. 188–221). Cambridge, England: Cambridge University Press.

Riazi, M., Shi, L., & Haggerty, J. (2018). Analysis of the empirical research in the journal of second language writing at its 25th year (1992–2016). *Journal of Second Language Writing*, *41*(7), 41–54.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387.

Rolheiser, C. (1998). *Self-evaluation: Helping students get better at it! A teacher's resource book.* Melbourne, Australia: Harper Brownlow Education.

Romano, F. (2019). Grammatical accuracy in EAP writing. *Journal of English for Academic Purposes*, *41*, 1–10.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, *11*(10), 1–13.

Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in grade 5-6 mathematics effects on problem-solving achievement. *Educational Assessment*, *8*(1), 43–58.

Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1999). Effects of self-evaluation training on narrative writing. *Assessing Writing*, *6*(1), 107–132.

Ross, S. J. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics*, *26*(3), 317–342.

Ruegg, R. (2018). The effect of peer and teacher feedback on changes in EFL students'writing self-efficacy. *Language Learning Journal*, *46*(2), 87–102.

Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*, *30*(3), 231–240.

Sadek, N. (2018). The effect of self-assessment as a revision technique on Egyptian EFL students' expository essay writing. In A. Ahmed & H. Abouabdelkader (Eds.), *Assessing EFL writing in the 21st century Arab world: Revealing the unknown* (pp. 21–52). Cham, Switzerland: Palgrave Macmillan.

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Sambell, K., McDowell, L., & Montgomery, C. (2013). Developing students as self-assessors and effective lifelong learners. In *Assessment for learning in higher education* (pp. 120–146). New York, NY: Routledge.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., King, J., & . (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*(6), 323–337.

Schunk, D. H., & Pajares, F. (2010). Self-efficacy beliefs. In I. M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *International encyclopedia of education* (pp. 668–672). San Diego, CA: Academic.

Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in Education. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 34-54).

Schunk, D. H., & Usher, E. L. (2011). Assessing self-efficacy for self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 282–297). London, England: Routledge.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence modeling, goal

setting, and self-evaluation. *Reading & Writing Quarterly*, *19*(2), 159–172.

Schunk, D. H, & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 631–649). Cambridge, MA: Academic Press.

Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology*, *81*(1), 91–100.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 4–14.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, *27*(4), 657–677.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*(3), 185–211.

Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Chanllenge and change in language teaching* (pp. 17–30). Bath, England: Heinemann.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510–532.

Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, *1*, 293–319.

Smagorinsky, P. (2008). The method section as conceptual epicenter in constructing social science research reports. *Written Communication*, *25*(3), 389–411.

Stefani, L. A. J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education*, *19*(1), 69–75.

Sullivan, K., & Hall, C. (1997). Introducing students to self-assessment. *Assessment & Evaluation in Higher Education*, *22*(3), 289–305.

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, *90*, 1–17. https://doi.org/10.1016/j.system.2020.102221

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of

work. *Higher Education*, *76*(3), 467–481.

Tan, K. (2007). Conceptions of self-assessment: What is needed for long-term learning? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the long term* (pp. 114–127). London, England: Routledge.

Tan, K. H. K. (2004). Does student self-assessment empower or discipline students? *Assessment & Evaluation in Higher Education*, *29*(6), 651–662.

Tan, K. H. K. (2009). Meanings and practices of power in academics' conceptions of student self-assessment. *Teaching in Higher Education*, *14*(4), 361–373.

Tang, F., & Xu, J. (2011). A study and survey of Chinese non-English majors' English writing self-efficacy. *Foreign Language World*, *6*, 22–30.

Teng, L. S., Sun, P. P., & Xu, L. (2018). Conceptualizing writing self-efficacy in English as a foreign language contexts: Scale validation through structural equation modeling. *TESOL Quarterly*, *52*(4), 911–942.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55–87). The Netherlands: Kluwer Academic.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294.

Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*, *38*(3), 323–342.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796.

Van Reybroeck, M., Penneman, J., Vidick, C., & Galand, B. (2017). Progressive treatment and self-assessment: Effects on students' automatisation of grammatical spelling and self-efficacy beliefs. *Reading & Writing*, *30*(9), 1965–1985.

Villalón, R., Mateos, M., & Cuevas, I. (2015). High school boys'and girls' writing conceptions and writing self-efficacy beliefs: What is their role in writing performance? *Educational Psychology*, *35*(6), 653–674.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, England: MIT Press

Wang, C., & Kim, D.H. (2017). English-language learners in Chinese high schools: Self-efficacy profiles. In C. Michael. & W. Yip (Eds.), *Cognition, metacognition and academic performance: An East Asian perspective* (pp. 27–39). New York, NY: Taylor and Francis.

Wang, C., Kim, D. H., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System*, *44*(1), 24–33.

Wang, W. (2016). Using rubrics in student self-assessment: Student perceptions in the English as a foreign language writing context. *Assessment & Evaluation in Higher Education*, 1–13.

Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: The crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education*, *43*(7), 1032–1047.

Weigle, S. C. (2002). *Assessing writing*. Cambridge, England: Cambridge University Press.

White, E. M. (1994). *Teaching and assessing writing: Recent advances in understanding, evaluating, and improving student performance* (2nd ed.). San Francisco, CA: Jossey-Bass Publishers.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accurcy and complexity (Technical report)*. Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Centre.

Wong, H. M. (2016). I can assess myself: Singaporean primary students' and teachers' perceptions of students' self-assessment ability. *Education*, *44*(4), 442–457.

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, *39*(4), 510–522.

Wright, T. (2006). Managing classroom life. In G. Simon & I. K. Miller (Eds.), *Understanding the language classroom* (pp. 64–88). New York, NY: Palgrave Macmillan.

Xiao, G., & Chen, X. (2018). Application of COCA in EFL writing instruction at the tertiary level in China. *International Journal of Emerging Technologies in*

Learning, *13*(9), 160–173.

Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports students' self-regulation in English language learning. *System*, *81*, 39–49.

Xu, Y. (2019). Scaffolding students' self-assessment of their english essays with annotated samples: A mixed-methods study. *Chinese Journal of Applied Linguistics*, *42*(4), 503–526.

Yan, Z. (2018). Student self-assessment practices: The role of gender, school level and goal orientation. *Assessment in Education: Principles, Policy and Practice*, *25*(2), 183–199.

Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, *45*(2), 224–238.

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, *42*(8), 1247–1262.

Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2019). Student self-assessment: Why do they do it? *Educational Psychology*, 1–20.

Yang, L., & Gao, S. (2013). Beliefs and practices of Chinese university teachers in EFL writing instruction. *Language, Culture and Curriculum*, *26*(2), 128–145.

Yin, R. K. (2018). *Case study research and application: Design and methods* (6th ed.). Los Angeles, CA: SAGE Publications.

Yoon, E., & Lee, H. K. (2013). Do effects of self-assessment differ by L2 language level? A case of Korean learners of English. *Asia-Pacific Education Researcher*, *22*(4), 731–739.

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., & Hu, G. (2017). Understanding university students' peer feedback practices in EFL writing: Insights from a case study. *Assessing Writing*, *33*(8), 25–35.

Zhang, J. (2016). Same text different processing? Exploring how raters' cognitive and meta-cognitive strategies influence rating accuracy in essay scoring. *Assessing Writing*, *27*, 37–53.

Zhang, L. J. (2013). Second language writing as and for second language learning. *Journal of Second Language Writing*, *22*(4), 446–447.

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. (张军，程晓龙). (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]. 西安外国语大学学报 *[Journal of Xi'an International Studies University], 28*(1), 48–56.

Zhang, L. J., & Zhang, D. (2018). Metacognition in TESOL: Theory and practice. In J. Liontas & A. Shehadeh (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 682–792). Malden, MA: Wiley.

Zheng, H., Huang, J., & Chen, Y. (2007). Effects of self-assessment training on Chinese students' performance on college English writing tests. *Polyglossia*, *23*, 33–42.

Zheng, X., & Adamson, B. (2003). The pedagogy of a secondary school teacher of English in the People's Republic of China: Challenging the stereotypes. *RELC Journal*, *34*(3), 323–337.

Zheng, Y., & Yu, S. (2019). What has been assessed in writing and how? Empirical evidence from Assessing Writing (2000–2018). *Assessing Writing*, *42*, 1–11.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183.

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). London, England: Routledge.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135–147.

Zimmerman, B. J, & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, *31*(4), 845–862.

Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The nature of problem solving* (pp. 233–262). New York, NY: Routledge.

Zimmerman, B. J., & Schunk, D. H. (2008). Motivation: An essential dimension of selfregulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 1–30). Mahwah, NJ: Erlbaum.

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). London, England: Routledge