Diagnostic Assessment of Discourse Competence in EFL Learners' Academic Writing in University Study

by

WANG Yumin

A Thesis Submitted to

The Education University of Hong Kong

in Partial Fulfilment of the Requirement for

the Degree of Doctor of Philosophy

August 2021

Statement of Originality

I, WANG Yumin, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the University's policies and regulations on Academic Honesty, Copyright and Plagiarism in writing the thesis and no material in this thesis has been submitted for a degree in this or other universities.

Abstract

Diagnostic assessment (DA) has its objective in generating fine-grained diagnostic profiles for effective treatment or intervention. Most existing DA studies on English academic writing focus on developing diagnostic tools to identify problems or difficulties within written texts. Rarely are there studies that design and implement a treatment or intervention phrase following the diagnosis. Moreover, despite of its importance, few studies have examined the construct of discourse competence. This thesis project, therefore, firstly diagnosed EFL learners' development of discourse competence (DC) as indicated by their academic writing in university, and then based on the diagnostic results, designed and implemented an intervention study.

The project involved six Chinese EFL learners studying in a university in Hong Kong, nine individuals studying in mainland, and their 277 peers pursuing their first degree in an international degree program based on a comprehensive university in mainland China. In both contexts, the participants need to write academic assignments in English. Methodologically, the project adopted a mixed-method approach and was conducted via four inter-linked studies, namely, a text analysis study, a comparative multiple-case studies, a survey, and an intervention study. The diagnosis phrase, consisting of text analysis, multiple case study and a survey, examined student writers' textual features, composing strategies, and academic writing knowledge status. The intervention phrase was designed to resolve students' problems and difficulties identified in the first phase.

The results showed that the cyclical diagnostic assessment procedure was effective in identifying specific lacks, problems and difficulties in learners' written products, composing process and knowledge status, and the diagnosis could inform the design of intervention

measures in teaching and learning of EFL academic writing. Specifically, contextuality was found to be a critical element for ensuring effective treatment following diagnosis. To secure quality diagnosis, however, proper teacher training or professional development program should be in place. The fine-grained rubric developed and validated in this project to assess text features at the discourse level would be useful for this endeavor. The rubric could also be employed by EFL writers for self-diagnosis. Furthermore, the project also developed a set of discipline-situated metacognitive intervention strategies that teachers could use to improve students' development of DC in academic writing. These outcomes of the project could enrich current understanding of construct specificity in diagnostic assessment.

Overall, through investigating discourse competence, an under-researched construct within the area of language assessment, and combining the diagnosis with an intervention study, this thesis project shed new lights on the design, development and implementation of diagnostic assessment tools in English for Academic writing purposes classrooms.

Keywords: diagnostic assessment; discourse competence; academic writing; EFL learners; university study

Acknowledgments

I have been conceiving this section since I embarked on my PhD journey five years ago. Over these years, I feel more than lucky to have precious guidance, help, support, encouragement and companionship from a number of people who have lightened me up on the sometimes endless and lonely journey. They are: my supervision team, my thesis exam panel, my school mates, my office mates, my flat mates, my friends, my students, my family, and more people in Hong Kong, mainland China and other parts of the world.

I own special debt of gratitude to my principal supervisor Dr. Xie Qin. She opened the door for me to have systemic exploration of diagnostic assessment of discourse competence in L2 writing, guided me through the whole research process and also provided part of the financial support for my study. I thank her for having belief in my abilities and having willingness to involve me in her project. Her ongoing suggestions and numerous provisions of timely and meticulous feedback have been invaluable in facilitating my research progress and enlightening my study ahead. I am also deeply infected by her vigorous commitment and rigorous manner in pursuit of academic research. Without her continuous push, patience and perseverance, my thesis study would be much less finished.

I also owe a great debt of thanks to my associate supervisors Professor Cheung Hin Tat and Dr. Paul Stapleton. I am grateful to Professor Cheung for his careful guidance in helping enhance the logic among the series of my studies and for his inspiring sharing over the lunch table addressing my confusion on some general humanity and social science issues. I also thank him for his simple but insightful suggestion of the chicken soup menu to release me from worries on the way of research. I am grateful to Dr. Stapleton for his systemic lectures and research work in stimulating my critical thinking and studies in L2 writing. I also thank him for his encouragement and timely suggestions on my research progress and his

considerate recommendation of a series of research articles and two special booklists to strengthen my qualitative research techniques and improve my logical thinking.

I am also extremely grateful to my thesis exam panel for their generous advice, meticulous comments, and kind encouragement. They include Prof. Constant Leung in King' College London, Dr. Cecilia Guanfang Zhao in University of Macau, Dr. Wang Lixun and Prof. John Trench in EdUHK. Reading their meticulous comments have made me relive all of my learning journey behind and beyond. Their considerate guidance and personal charisma have shed more than light and inspiration for my academic journey ahead.

I am also deeply indebted to the professors and teachers who stimulated my doctoral dream, offered kind suggestions, and provided warm encouragement. They include Dr. Jiang Hongliu in Sichuan University, Prof. Guan Hefeng in Henan University, Prof. Gu Xiangdong in Chongqing University, Prof. Paul Kei Matsuda in Arazona State University, and Prof. Lawrence Jun Zhang in University of Auckland. Particularly thanks are due to Dr. Jiang Hongliu for her constant care and support over the years, which stimulated me to pursue my doctoral study and live up to my own potential in both life and work.

Thanks are also due to the administrative staff and my colleagues both in EdUHK and in Henan University in mainland who provided immediate support and help in research preparation and data collection procedures. They include Ms. Grace Lau Ho Yan and Ms. Sarah Ng Oi Ting Sarah in Graduate School in EdUHK, and in mainland China Prof. Zhao Xinshun, Prof. Ma Xiaoyong, Mr. Kang Tianfeng, Ms. Niu Xian, Mr. Du Haonian, Ms. Su Fangfang, Ms. Zhao Shuhong, Ms. Miao Lili, Dr. Chen Xiaoyan, Ms. Li Xiaoxi, Ms, Wang Jiashi, Ms. Qin Xiao, Ms. Du Aiping, Ms. Niu Yongjuan, Ms. Wang Lijuan, Ms. Wang Aiyu, Ms Ma Di, Ms. Hu Feng, Ms. Guo Jinyu, Ms. Yang Junqing, Mr. Li Shuchen, and still many others beyond my naming here. Particular thanks are to Ms. Wang Zihan, Ms. Jia Lijuan, and Ms. Zhao Jingxuan for their timely support and careful commitment to the rating work.

I also thank the more than 15 participants and 300 students both in my pilot study and in the full thesis project. Thank them for their confidence in me and their willingness to spare their time and share their writing stories with me. Some of their constant contact with me over the two years after my study with them has further strengthened my confidence and determination in carrying on the project forward.

I am also grateful for the companionship of and support from my schoolmates and friends in EdUHK and in mainland. They provide constant nutrition physically, mentally and spiritually. They include Dr. Miranda Lan, Dr. Chen Shimin Christy, Dr. Yang Hongyu, Dr. Sun Liying, Dr. Yang Shuxia, Dr. Wu Yi, Dr. Li Yulong, Dr. Yang Huaxin Peanut, Dr. Zhi Yuying Cathy, Dr. Dong Lina, Dr. Hu Xiaojuan, Dr. Chai Yemao CATHY, Dr. Chen Xi, Dr. Zheng Qun, Dr. William Huang, To-be-Dr. Wu Xue Alice, To-be-Dr. Zhang Huanhuan, and Ms. Wang Mengjuan. Particularly, Dr. Miranda showed me the charm of research and taught me the value of life with her more than 60 years life stories. Peanut walked with me almost hundreds of times around the trails up and down the campus and shared her research and life stories all through these years. Yulong helped strengthen my research awareness with his sparkling research experience and skills. Xiaojuan called me tens of times for playing badminton and CATHY taught me swimming to relieve my frustration in study. Christy brought self-cooked soup to my worktable in the very early morning. Lina brought me the mooncake when I was lost in piles of books in the Mid-Autumn Festival. Cathy helped pay my residence bill when I missed the date. Alice asked me to go hiking when she perceived my obsession with reading literature during the weekend. So on and so forth.

Finally, thanks are due to my family. Thank my parents for always reminding me to take good care of myself and saying over the phone they are well even if my father had been ill in hospital for two weeks. 'Have you got your degree?' Father used to ask me this question upon seeing me or hearing me over the phone. I wish he could hear my Yes and smile from the sky.

Thank my elder brother who has been taking care of my parents and also giving me constant mental comfort as best as he can. Thank my husband Dr. Su for his technical, mental and domestic support in my doctoral study. I am also deeply indebted to my daughter for my disappearance from most of her primary school activities. Thanks for Cheng's accompany. My thesis study would not be possible without their love, support, patience and understanding.

Thank you all! You have raised me up more than I can be!

Table of Contents

List of Abbreviations

AL	Academic literacies
AMOS	Analysis of a moment structures
ANOVA	Analysis of variance
AWE	Automatic writing evaluation
CBI	Content-based instruction
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CLIL	Content and language integrated learning
CG	Control group
DA	Diagnostic assessment
DC	Discourse competence
DF	Discourse feature
DO	Discourse overall
DS	Disciplinary knowledge
EAP	English for Academic Purpose
EFA	Exploratory factor analysis
EFL	English as a foreign language
EG	Experimental group
EMI	English as a medium of instruction
ESP	English for specific purposes
ESAP writing	English for Specific Academic Purpose writing
EVL	Evaluation
FML	Formulation

GC	Global coherence
GE	General education
GN	Genre knowledge
HP	High proficiency
IBM	International Business Machines Corporation
IC	Intercultural communication
IELTS	International English Language Testing System
KO	The overall academic writing knowledge status
L1	First language
L2	Second language
LA	Language skills
LP	Low proficiency
LW	Learn to write
MFRM	Many-facet Rasch Measurement
MP	Medium proficiency
NNS	Non-native speakers
PCA	Pcomponent analysis
PLN	Planning
RG	Reasoning and argumentation
RMSEA	Root mean square error of approximation
RQ	Research question
RSC	Research
RVS	Revision
SEN	Special education needs
SFL	Systemic functional linguistics

SPSS	Statistical Package for the Social Sciences
SRMR	Standardized Root Mean Squared Error
TERA	Text Ease and Readability Assessor
TLI	Tucker–Lewis index
TOEFL	Test of English as a foreign language
WL	Write to learn
WLC	Writing to learn content
WLL	Writing to learn language

List of Figures

List of Graphs

List of Tables

Chapter 1 Introduction

Discourse competence (DC) is a critical component in writing as it highly influences coherence, cohesion and overall flow of the text. In language learning and teaching research, there has been extensive studies on DC in L1 context. However, in the L2 context, there is rare explicit discussion on the exercise of this competence in classroom learning. This study conducted a systemic diagnostic assessment of EFL writers' problems and difficulties in developing DC in English academic writing. Specifically, three tiers of diagnosis were conducted: the first tier on written scripts, the second on the processing strategies, and the third on academic writing knowledge status. Multiple sources of data are collected centering on a reading-integrated essay writing task including students' written scripts, semi-structured interviews, writer logs, surveys, and intervention data. The following introduction mainly serves to set the scene for the theoretical understanding of DC, a snapshot of the research context, an overview of the research design, and an introduction to the general structure of the thesis.

1.1 Discourse competence in English academic writing

Within language studies, the term ‘discourse’ has been used in multiple ways. Some linguists (e.g. Halliday & Hassan, 1976; Leech, 1983) considered ‘discourse’ similar to ‘text’, both are used to refer to written and spoken language; some scholars (e.g., Widdowson, 2007) kept ‘discourse’ for written language, and still some (e.g. Coulthard & Condlin, 2014) preferred to use ‘discourse’ for spoken language and ‘text’ for written language. One consequence of the multiplication of interpretations of ‘discourse’ is that there are multiple terms associated with discourse features, such as cohesion, coherence, structure, organization, rhetorical patterns, style, register, and genre. The present study followed Halliday and Hassan (1976) and considered ‘discourse’ as referring to both written and spoken language.

Discourse competence in writing can be understood as the knowledge and skills of managing linguistic features and their semantic function within a specific text and/or context beyond the word and sentence level form and functions (Canale & Swan, 1980; Harris, 1952; Hymes, 1972; Paltridge, 2012). It contributes to both linguistic accuracy and social appropriateness of the written discourse (Canale & Swan, 1980). There are frequently different perspectives on

the conceptualization of concrete linguistic features of DC due to different research purposes. For instance, with a particular interest on ethnography of spoken communication, Hymes (1972) conceived discourse knowledge as consisting of not only syntax but also cultural components by the integration of linguistic and anthropological perspectives. With specific focus on writing, Grabe and Kaplan (1996) considered a cohesive text as structured by recognizing main topics, organizing scheme, transition markers, informational structure and semantic relations across clauses. Despite various scope and approaches to 'discourse' competence or discourse level qualities in writing, DC is widely acknowledged to be an essential component of communicative competence which is central to the mastery of academic writing (Bachman, 1990; Bruce, 2008; Canale & Swain, 1980; Evans & Morrison, 2011; Hyland, 2011).

1.2 The context of English for Academic Purpose writing in university study in the L2/EFL context

In English medium degree study program, there are two main strands of English for Academic Purpose **(**EAP) writing, namely writing to enhance academic literacy (similar to the 'learn to write' concept (Wingate, 2018; Xu,2015) and writing to facilitate disciplinary

studies across courses (similar to the ‘write to learn the content’ concept) (Flowerdew, 2020; Gardner, 2012;). Disciplinary writing is crucial to the achievement of academic success (Nesi & Gardner, 2012). Students are required to write course papers, research proposals, and degree thesis in English. Development of effective academic writing skills is thus an essential objective in university writing curriculum. Bruce (2008) pointed out that to promote English academic learners’ writing skills, the primary focus on atomized units of language at sentence or clause level is not only inadequate but may also be misleading. The genre-based pedagogy gives particular emphasis on the role of text and context in writing (Hyland, 2003). Bruce (2008) proposed the genre-based approach to promote the development of DC. But the focus of most existing studies is primarily on English academic learners situated in L1 context. It may be meaningful to examine the exercise of EAP writing in L2 or EFL context.

With the trend of internationalization in higher education and the role of English as an international lingua franca, more and more EFL students are undertaking EAP writing in universities outside of English-speaking countries. In China, the context of the present research, the need of EAP writing in Sino-foreign cooperation programs is rather similar to

those in EMI universities in English speaking countries. The EAP writing course in this context and its like usually serve three purposes: to develop L2 proficiency, to facilitate disciplinary studies and to enhance academic literacy. In the first case, academic writing in English is focused on facilitating the learning of English, that is writing is considered to be a vehicle or means, while developing language proficiency is the end (Manchón, 2011). In line with this purpose, teaching practice often involves in form-focused instruction, provision of feedback and exploration of learner motivations and beliefs. The second purpose of EAP writing, namely, 'writing to learn the content', on the other hand, requires the support provided by both writing and content/subject teachers (Hirvela, 2011). The third purpose of learning to write has been investigated from three perspectives: the text (Hyland, 2007), the writers (Hyland, 2005a) and the readers (Hyland, 2005b). Pedagogically, teachers of writing need to understand three core components: the learners, the texts that learners will need, and the contexts in which the writing activities are likely to happen (Hyland, 2015).

In the context of mainland China, prior to university study, a large portion of students' English learning experience involve learning discrete linguistic items (such as lexis and

syntax) and the practice of timed short compositions for exam purpose, often in 100-250 words. Writing in English is primarily targeting the first purpose aforementioned. This traditional approach, however, cannot fulfil the needs of joint educational programs at the university level, where English writing has to meet future professional needs as well as academic study. Unsurprisingly, therefore, students coming out of the school system often encounter challenges when they entered the university and are required to write extended pieces in English for their disciplinary work.

1.3 Diagnostic assessment for language learning and teaching

Upon entering the university, these Chinese students have accumulated certain knowledge of vocabulary and grammar of English, but such knowledge remains insufficient for them to cope with the English writing tasks in university (Gao, 2007; Shi et al., 2019). It is necessary to have more specific understanding of these learners' writing strengths and weaknesses for better future improvement. Diagnostic assessment, i.e., the act of precisely analyzing a problem and identifying its causes for the purpose of effective treatment (Rupp et al., 2010), arises as a helpful approach to cope with this situation.

Existing research on diagnostic language assessment has identified some key principles, systemic procedures, and a prolific source of tools to help generate a fine-grained profile of language learner problems and difficulties (Harding et al., 2015; Knoch, 2009; Rupp et al., 2010; Xie, 2020). For instance, Harding et al. (2015) proposed a set of five principles for diagnostic language assessment. Their study has provided a systemic guide in preparation for a well-designed, valid and meaningful diagnostic work in the context of language assessment. Knoch (2009) developed and validated a diagnostic tool for assessing writing abilities of first year university students in a New Zealand university. Llosa et al. (2011) identified the most prevalent and challenging types of writing tasks learners face in the context of secondary school study in America; they also investigated learners' challenges in the composing processes. Fox and Artemeva (2017) developed a diagnostic writing task and a corresponding rubric aimed at identifying undergraduate students academic writing needs in engineering study in a Canadian university. Xie (2020) adopted an item bank approach and developed a set of diagnostic test to generate specific and precise information about linguistic problems in EFL university student writing.

While existing research on diagnostic language assessment has provided multiple inspirations and insightful guidance, still relatively few empirical diagnostic assessment works are conducted in English academic writing, of which most studies up to now have mainly examined learner problems/difficulties in particular linguistic categories or on a general writing model, through textual analysis or verbal report. That is, most diagnostic assessment work on student writing focus on diagnosing textual issues. As pointed in some studies (Zhao, 2019b), L2 writing development and difficulties may be invisible in the written products, and students may not be able to articulate precisely their own writing difficulties. While remediating a problem at the place where it appears may only work for some time, but it does not last long and may not be transferred into a different context. Therefore, it is necessary to probe into students' discourse level features and problems/difficulties from multiple perspectives and at different levels so as to enable the provision of effective treatment in facilitating the students' development of DC overtime and across different fields of study.

1.4 Aim and significance of the present study

1.4.1 Aim

The primary goal of the thesis study is to conduct a systematic diagnostic assessment in the expectation of identifying specific problems and difficulties of Chinese EFL undergraduate students in essay writing across the curriculum and help facilitate their development of discourse competence at the early stage of their university academic writing journey. Considering the time and space restraints, the present project mainly focused on discourse features rather than discourse skills and knowledge under the overarching concept of DC. Specifically, the present study conducted a three-tiered diagnostic assessment of student writers' textual features, composing strategies, and knowledge about academic writing to identify their weaknesses in writing and to investigate potential influencing factors. The focal participants were a group of Chinese EFL undergraduates majoring in business studies; and mixed-methods were adopted to collect data including text analyses, comparative case study, survey, and an intervention study. Text analyses of students written scripts were conducted to validate the diagnostic rubrics designed to assess the written texts produced by the participants. Drawing on analysis of written scripts, writer logs, and semi-structured interviews, EFL learners' discourse-level weaknesses were examined together with their writing strategies and academic writing knowledge status to ascertain their relationships.

1.4.2 Significance

The present study has potential to contribute to theory and practice due to the importance of the DC construct, the research method, and the uniqueness of the research context.

First, DC is a crucial component in academic writing while writing is a complex process that encompasses dynamic interactions of diverse components in ever-changing social and educational contexts. In view of centrality of DC and complexity of writing, there is a need to keep examining the traits of DC in writing in a timely manner.

Second, the proposed three-tiered diagnostic procedure is a comprehensive approach for exploring specific problems and difficulties in student learning. It is innovative in relating the textual features to the composing process and the underlying knowledge status to create a rich insight into particular writing difficulties. The main purpose of diagnostic language assessment is to generate a fine-grained diagnostic profile of learners' specific problems and weaknesses in language learning in the expectation of provision of effective treatment or intervention measures. Regarding student writing practice, most studies in language testing

and assessment filed focus primarily on students written scripts, with little mention to the actual composing processes or exploration of pertinent influencing factors. Still many existing studies only addressed features at one level, be it the textual level or the process level, but not the complete cycle of the writing activity itself. The present study intends to relate students written scripts to their composing processes and academic writing knowledge status to uncover students' specific problems and difficulties in English writing. This three-tiered diagnostic procedure is beneficial to our understanding of the multifaceted nature of L2 writing and also helpful for our understanding of the pertinent factors and processes and the provision of effective treatment measures.

Thirdly, the project is of great value to research into DC in academic writing in the EFL context. DC is a crucial component in effective writing, yet there is inadequate research attention to this topic in research on EFL teaching and learning practice. This study focuses on the genuine exercise of novice EFL writers' development of DC in a Chinese context. The findings may shed new light on understanding EFL learners' acquisition of DC in language

learning. It may also provide useful implications for teaching and learning of English academic writing in many similar EFL/L2 learning contexts around the world.

1.5 Research context

The present study is situated in the context of an international joint degree program (based in China). Compared to conventional higher education in China and English-medium instruction universities overseas, this program has its special characteristics regarding human resources, teaching and learning materials, institutional expectations and political contexts.

In respects to human resources, the teaching staff in this joint education program consists of foreign language teachers, local language teachers, foreign subject teachers, and local subject teachers. They share offices and meeting rooms and venues of teaching. A joint management committee consisting of members from both local and overseas sides is also organized to be responsible for daily operation issues. There are some annual teaching and learning symposium activities across the program. While in conventional university programs in mainland China, language teachers and subject teachers are usually dispersed in different

departments or faculties; and there are rare opportunities for language teachers and subject teachers to exchange professional views of their students.

Regarding the teaching and learning materials, in the context of the joint international educational program, many textbooks and teaching materials are introduced from abroad. To meet curriculum needs and achieve academic excellence, students are expected to compose academic essays in English both for their language courses and disciplinary courses. In other words, compared to conventional tertiary education in China, students in the joint educational program have more needs in English writing and hence more challenges and opportunities to develop their English writing proficiency. Consequently, there is intense writing-related exercise in this program, including a special writing course offered by foreign language teachers, complementary writing lectures in the reading/writing course by local language teachers, and intermittent instructions from different subject teachers as well.

Regarding institutional expectations and political context, students in this joint program not only need to satisfy curricular expectation in home university but also have to achieve the

objectives required in the partner university to complete their degree study. For instance, students need to attend the compulsory courses of *The Principles of Maxism* and *Mao Zedong Thoughts* and *Morality and Law*; meanwhile, they also need to complete a series of compulsory disciplinary courses regarding international business studies. Hence, there is a very tight teaching and learning schedule in this bilingual instruction context.

Altogether, the abundant disciplinary writing exercises in the research site provide a meaningful venue for the present study to gain detailed insights into EFL learners' genuine academic writing situation; the collaborative environment between language teachers and content lecturers assures the availability of valuable source for later remedial action. Meanwhile, the intense teaching and learning schedule also poses potential challenges for subsequent design, development and implementation of related intervention writing activities.

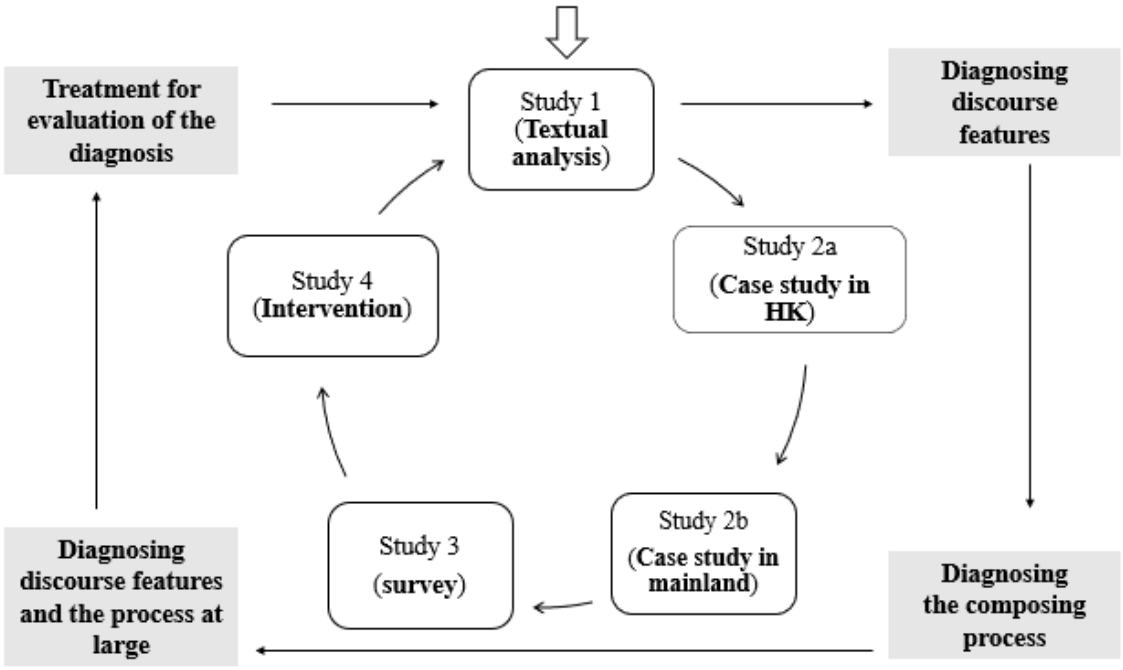
1.6 Overall research design

The present study encompasses three core components in diagnosis assessment, i.e., diagnosis, intervention and evaluation, as schematized in Figure 1.1 below. The research was conducted in three phases. First, individual case studies were conducted aiming at examining

students' particular discourse level features and their composing processes to gain insights into the underlying composing strategies that may affect discourse-level performance in the written product. As suggested in classical linguistic studies (He, 2003; Lv, 1977; Zhang & Chen, 1981), comparison is an essential method in understanding a phenomenon or subject as only by comparison can we gain a relatively complete picture of our target subject based on necessary reference or parameter. To enable comparisons, case studies were conducted both in Hong Kong and in the mainland China. Within each site, individual essay writing processes across the curriculum were followed over one semester. The findings from individual case study formed the foundation for the later large-scale survey study and a subsequent intervention study. At the 2nd phase, a large-scale survey study was conducted to examine students' processing strategies at a wider scope and check the generalizability of the findings from case studies.

Figure 1.1

A systemic and cyclical diagnostic assessment design



Both the case study and the survey study informed the design of an intervention study at the third phase. The intervention study aimed at addressing the discourse problems and difficulties as identified in students' written essays and composing processes. To evaluate the effectiveness of the intervention program, analysis of pre-test and post-test written essays and semi-structured individual interview were also conducted. Effective intervention in turn provided useful evidence for justifying the diagnosis results.

1.7 Thesis structure

This thesis documented the above research in eight chapters. Chapter 1, the current one, sets the scene and presents a brief overview of the research design and thesis structure. The following four chapters, Chapter 2, 3, 4, and 5, present a series of empirical studies on diagnosis, intervention and evaluation of the research. Chapter 2 focuses on a text-based study, Chapter 3 on a comparative multiple case study in Hong Kong and in mainland China, Chapter 4 on a survey study, and Chapter 5 on an intervention study. Each of the four main chapters provides a focused literature review of the key concepts involved in each study, explains the details of the research design, clarifies the rationale for the decisions made on data collection and analysis methods, reports and discusses the findings. One major benefit of this organization strategy is that each study could be presented more holistically with a clear focus of its own; this arrangement, resembling the portfolio approach, allows the present researcher to publish the research from the thesis directly.

Specifically, Chapter 2 reports a study focusing on the quantitative analysis of 108 undergraduate students written scripts. Chapter 3 reports a comparative multiple case study of six participants in the Hong Kong context and nine participants in the mainland context.

Chapter 4 reports a survey study of 277 students in the mainland context. Chapter 5 reports an intervention study conducted in the mainland context. Chapter 6 synthesizes the findings reported in the previous chapters and discussed them in relation to the development of discourse competencies in L2 undergraduate student academic writing. Implications were drawn to inform curriculum and syllabus design and teaching of an ESAP writing course in international programs at the tertiary level in China. Chapter 7 ends the thesis with an overall conclusion and limitations and directions for future study.

Chapter 2 Study One — A Textual Study

Chapter Abstract

Discourse competence is a critical component in writing as it influences coherence, cohesion and the overall flow of the text. This study reports on the design and development of a set of discourse assessment rubrics for diagnosis of EFL writers' problems and difficulties in their academic writing exercise. A sample of 108 disciplinary written scripts were selected from students studying in an international joint degree program at a university in mainland China. The study found that the students were stronger in *logical connectives* but weaker in *topic building*, *global coherence* and *complexity of hedges and boosters*. The findings enriched work on construct specificity in diagnostic assessment research and also shed light on practical pedagogy for classroom instruction in EFL/L2 writing course.

2.1 Introduction

Discourse competence (DC) is a critical component in writing as it influences coherence, cohesion and overall flow of the text. Extensive studies on DC have been conducted in the L1 context (Miller & Pessoa, 2016; Tardy & Swales, 2009); there is rare explicit discussion in

the L2 context on the exercise of this competence, especially in classroom learning. This study designed and developed a set of discourse assessment rubrics and explored their application in diagnosing EFL writers' problems and difficulties in their written essays.

The following section will first survey existing literature to develop an operational understanding of DC in assessing writing, and then examine existing diagnostic assessment instruments and strategies to discern discourse-related qualities in writing. Finally, it reviews empirical studies on EFL learners' writing problems and difficulties in university study.

2.2 Literature review

2.2.1 DC in EFL writing

In general, DC in writing can be understood as the knowledge and skills of managing linguistic features and their semantic functions within a specific text or context beyond forms and functions at the word and sentence levels (Canale & Swan, 1980; Harris, 1952; Hymes, 1972; Paltridge, 2012). DC contributes to both linguistic accuracy and social appropriateness of the written discourse (Canale & Swan, 1980). However, the operation of DC in empirical studies vary; different linguistic features were adopted to denote DC to serve different

research purposes. At the early stage, with a particular interest in ethnography of spoken communication, Hymes (1967, 1972) proposed the concept of 'communicative competence' to account for different forms of knowledge that contribute to an individual's competence to communicate effectively within a specific setting. By integrating linguistic and anthropological perspectives, Hymes (1972) conceived communicative competence as comprising linguistic and social-linguistic competence; discourse knowledge is part of the latter, comprising syntax and structure that convey particular cultural values and beliefs, such as participant age, social distance and the interaction between them (cf. Celce-Murcia, 2007).

In the fields of linguistics and applied linguistics, the most well-known models of communicative competence are the ones in Canale and Swain (1980), Bachman (1990) and Grabe and Kaplan (1996). Drawing on the theoretical descriptions of language use by Hymes (1972), Canale and Swain (1980) proposed a four-dimensional model of communicative competence to inform language teaching and learning: grammatical competence, DC, sociolinguistic competence and strategic competence. The first two dimensions refer to the use of the linguistic system itself, and the latter two define the functional aspects of

communication. In particular, they conceived DC as ‘mastery of how to combine and interpret meanings and forms to achieve a unified text in different modes by using (a) cohesion devices to relate forms and (b) coherence rules to organize meanings’ (Canale & Swain, 1980, p. 339). That is, in their model, DC emphasises unification of meaning and form through coherence rules and cohesive devices. However, as indicated by some researchers (Mauranan, 1996; Widdowson, 2001), this model only provides a list of components, without explaining the relations or interactions among them in actual performance.

Following Canale and Swain’s (1980) original framework, Bachman (1990) proposed a more refined model for guiding and informing language testing studies. Bachman (1990) proposed three dimensions of communicative language abilities from the communicative language perspective: language competence, strategic competence and psychophysiological mechanisms. The dimension of language competence further comprises organisational competence and pragmatic competence, with the former focusing on grammatical accuracy and the latter on appropriateness in context. Organisational competence can be further

divided into grammatical (vocabulary, syntax and phonology/graphology) and textual competence (comprising cohesion and rhetorical organisation). Cohesion in this framework comprises reference, substitution, ellipsis, conjunction and lexical cohesion, as well as conventions governing the ordering of old or new information in discourse. Rhetorical organisation pertains to the overall conceptual structure of a text (e.g., narration and description). This model has been a resource for test development in many contexts and represents the current state of the art (Purpura, 2008). However, as indicated by some authors (Knoch, 2011; Mauranan, 1996), the components of DC conceived in this model are rather limited, and thus, we need to find alternative ways of organising discourse and see whether and to what extent these are language-specific and can be described simply enough to be teachable and assessed in foreign language classes.

In the specific field of writing, Grabe and Kaplan (1996) proposed a model of writing as communicative language use. This model specifies communicative language use as comprising a context for language use and a representation of language users' verbal working memory. The verbal working memory has three subparts: internal goal setting, verbal

processing and internal processing output. Verbal processing again comprises three parts: language competence, knowledge of the world and online processing assembly. Language competence further comprises three competencies: linguistic, discourse and sociolinguistic. Their discourse knowledge refers to knowledge of the ways in which cohesive text is constructed, such as recognising the main topics, organising scheme, transition markers, informational structure and semantic relations across clauses. The Grabe and Kaplan (1996) model provides a way to integrate the three major concerns in developing a theory for writing: the writer's cognitive processing behaviour, linguistic and textual resources and contextual factors. Their organising scheme of linguistic and discourse competencies corresponds to Bachman's (1990) rhetorical organisation and that of transition markers and information structure corresponds to cohesion. However, their conceptualisation of discourse knowledge is especially related to writing and has more concrete specifications on recognising topic, information structure and semantic relations across clauses.

Another recent development in academic writing studies is related to the notion of voice or reader–writer interaction. It was suggested that the key to the concept of coherence is not

something that exists in words or structure, but how people make sense of what the reading or hearing (Yule, 2006). Hyland (2008) proposed an interactional model of voice in academic writing to understand how writers use language to plausibly represent an external reality as well as acknowledge, construct and negotiate social relations in building a convincing discourse. Hyland's interactional model comprises two sub-dimensions: writer-oriented stance and reader-oriented engagement. So far, most literature has conceived DC from the trait/text perspective; however, reader interpretation of the coded text is indispensable in the construction of information flow. Hence, it is beneficial to include the reader–writer interaction perspective in conceptualising the DC construct, as suggested by Chapelle (1998) and Chalhoub-Deville (2003).

The complex features of DC make it necessary to decompose this concept into sufficiently detailed components, to yield meaningful assessment information and trigger actionable teaching and learning support. Combining Canale and Swain's (1980) conceptualisation of DC in language teaching and learning, Bachman's (1990) textual competence in language assessment and Grabe and Kaplan's (1996) discourse knowledge in writing, the present study

operationalised DC as comprising five major components that are linguistically researchable and transferable across contexts: topic building, global coherence, local coherence, logical connectives and reader–writer interaction. Detailed definitions of these five constructs are shown in Table 2.1 below. Following Bachman (1990), we believe that the interactions between these various components and language use context characterise communicative language use.

Table 2.1

Operational conceptualisation of DC in English academic writing

No.	Dimension	Definition
1	Topic building	An authorial opinion or a main argument is established to suggest an interesting and meaningful direction.
2	Global coherence	Pertains to the overall conceptual structure of a text, conveying the rhetorical pattern characteristic of, and appropriate to, the target communicative purpose.
3	Local coherence	Mainly involves thematic progression patterns related to the distribution of given/new information between sentences.

4	Logical connectives	The use of conjunctions and adverbial adjuncts to show the logical or semantic relations between sentences and paragraphs.
5	Reader–writer interaction	Informs how writers position themselves and their readers through typical linguistic devices (such as self-pronouns, reader-pronouns, hedges and boosters).

2.2.2 Assessment of discourse-level qualities in EFL writing

This section examines existing instruments and strategies developed to assess the five discourse components as identified in the section above, especially in second and foreign language academic writing. Four main aspects of the assessment are reviewed: assessment tools, scope, purposes and target population. Based on a review of existing assessment instruments, corresponding strategies are proposed specifically for the present diagnostic purpose of EFL writing in university study.

The most commonly used assessment tool in a writing assessment is the rating scale or rubric. Furthermore, most assessment rubrics in writing have been developed with the main intention of assessing short essays (Brown & Abeywickrama, 2013; Hamp-Lyons, 1990;

Jacobs et al., 1981; Knoch, 2009), with little focus on extending academic writing in university study (Bruce & Hamp-Lyons, 2015). With the advancement of technology, automatic writing evaluation (AWE) and text analysis tools have become increasingly available to help identify students' writing features. Commercially available AWE tools are MY Access!^{TM} (by Vantage Learning), e-rater Criterion^{SM} (by Educational Testing Service, US), Intelligent Essay Assessor^{TM} (by Pearson Knowledge Technologies, US), Project Essay Grade (by Measurement Incorporated, US) and Pigai in China. Some web-based text analysis tools, such as Coh-Metrix, are also available (McNamara et al., 2013). Most of these automatic evaluation tools can provide both numerical scores and certain qualitative description and evaluative feedback with respect to vocabulary, sentence structure, organisation and style. However, for deep-level factors, only vague holistic positive comments are often provided, without indicating specific problems (Chapelle et al., 2015).

The development and validation of rubrics have been a constant focus in diagnostic assessment (DA) studies related to L2 writing (Knoch, 2009; Kim, 2011; Xie, 2017). Weigle (2002) suggested five aspects to be considered in the scale development process: (1) type of

rating scale to be desired, (2) principal users of the scoring rubric, (3) most important aspect(s) of writing and their sub-categorisation, (4) specific points or scoring magnitudes to be used and (5) ways in which scores are reported. First, regarding the rubric types, most existing DA studies have adopted analytical rubrics instead of holistic and primary trait rubrics (Fox & Artemeva, 2017; Knoch, 2009; Lockwood, 2016; Read, 2016). Analytical rubrics provide rich and detailed criteria to evaluate learners' specific strengths and weaknesses in several areas of writing; therefore, they provide more diagnostic information for individualised or tailored instruction/guidance. Second, regarding principal users, most rubrics have been developed as more rater-oriented and policy-maker-oriented rather than tester-oriented. Given its learning-oriented purpose, the primary users of DA rubrics are the learners themselves and their teachers. Hence, later development of rubrics for diagnostic purposes should consider assigning more agent roles to users instead of assessors or constructors. Third, regarding the target criteria to be included in a writing rubric, the target skills to be assessed and the exact weight assigned to each skill in DA are closely related to the instructional needs along with particular linguistic theories (Ferris & Hedgcock, 2014; Knoch, 2011). Fourth, in terms of scoring magnitudes, the following three key elements need

to be addressed: number of levels, justification for distinguishing between levels and descriptor formulation styles (Knoch, 2009). Particularly, the number of levels necessary for a rating scale is determined by the range of performances to be expected and the ways in which the test results are expected to be used (Bachman & Palmer, 2016; Knoch, 2011; Weigle, 2002). Rater background and experience can also influence the number of scaling points to be used. Bachman and Palmer (2016) recommended using more scaling points than the number of decisions to be made, because independent raters will not necessarily agree on exact scales. Furthermore, Harding et al. (2015) argued that DA should be suitable and efficient for administration in the classroom and generate rich and detailed feedback for the test takers. Many large-scale examinations use between six and nine steps. Finally, score reporting concerns the provision and reception of feedback. For diagnostic assessment to fulfil its purpose of effective treatment, students' take-up of the diagnostic feedback is a crucial stage (Hamp-Lyons, 2019; Knoch, 2009). That is, score-reporting in DA should be able to enhance the diagnosis-treatment interface.

In terms of the assessment scope, existing assessments on discourse-level qualities are mostly set within a wider perspective together with other language components (e.g., Bruce & Hamp-Lyons, 2015; Knoch, 2009; Xie & Lei, 2021). For instance, Xie and Lei (2021) diagnosed EFL learners' writing features in seven domains, but without explicitly stating the connection of related components in these domains to DC. Only a few studies have conducted a specific assessment of discrete discourse components, but in the form of a checklist (e.g., Kim, 2011), with no explicit systemic deconstruction to address DC as a coherent cluster of features. DC is an essential yet elusive concept in academic writing. In particular, for most novice EFL learners, it is a novel concept. Thus, this concept must be decomposed into sufficiently detailed diagnostic constructs for novice EFL learners to take up the DC concept and gradually acquire related abilities.

In terms of assessment purposes and target population, many writing rubrics have been published for high-stakes tests (such as those used in TOEFL and IELTS); some rubrics have been designed for diagnostic purposes on a large university scale (Knoch, 2009). However, only a few studies have focused on DA in the classroom context (Doe, 2011; Fox et al.,

2016). In particular, many universities have adopted diagnostic language assessment for addressing students' language needs to strengthen their engagement in disciplinary study. Such assessment is referred to as the Diagnostic Language Testing System in Europe (e.g. Alderson, 2005), Post-admission Assessment of Language in Australia (e.g., Knoch et al., 2016), Diagnostic English Language Needs Assessment in New Zealand (e.g., Read, 2008) and Diagnostic English Language Tracking Assessment in Hong Kong (e.g., Lockwood, 2013). However, the rubrics developed from these studies might not consider catering to the specific needs of particular groups of students in a particular syllabus. As a result, their application to the classroom teaching and learning context remains unexplored, as the teaching, learning and assessment context may vary considerably in terms of factors such as time arrangement, assessment resources and learner diversity. To address this issue, the present study explores the use of diagnostic rubrics for assessing EFL learners' specific strengths and weaknesses in DC in a classroom setting in the EFL context. Concerns about this learner group's special needs can further enrich the methods in the design and development of a diagnostic rubric.

Finally, in relation to assessment of the five discourse dimensions as identified in previous section, a close examination of existing assessment rubrics identified ten assessable features for practical diagnosis of novice L2 learners' writing. Specifically, within topic building, three sub-components were identified: topic relevance, thesis development and controlling idea (Brown & Barley, 1984; Bruce & Hamp-Lyons, 2015; Jacobs et al., 1981; Kim, 2019). Within global coherence, introduction, conclusion, and body parts were the specific concerns (Bruce & Hamp-Lyons, 2015; Kim, 2019; Lockwood, 2016). With local coherence, the typical concern was theme-rheme and Given/New information structure (Brown & Abeywickrama, 2013; Kim, 2019). With logical connectives, three features often discussed were range, accuracy, and adequacy (Knoch, 2009; Ong, 2011). Within reader-writer category, two typical features evaluated were frequencies and complexities of hedges (Knoch, 2009; Zhao, 2010) and boosters (Zhao, 2010). As a result, the ten specific discourse features were selected as the practical assessment categories in the present study.

2.2.3 Empirical studies on discourse-level problems and difficulties in EFL writing

The writing problems and difficulties of university EFL learners have been extensively studied from lexical (e.g., Chong et al., 2014; Xie, 2020), syntactical (e.g., Chong et al.,

2014; Xie, 2020) to discourse levels (e.g., Chong et al., 2014; Liang, 2006; Ong, 2011). Within the discourse level, both global (e.g., Liang, 2006) and local features (e.g., Ong, 2011) have been studied; synthesis studies have also been conducted (e.g., Hinkel, 2011). This section synthesizes existing empirical studies on written discourse features in relation to the five main categories and ten specific features as identified in the previous sections.

Regarding topic building, a frequent observation was that students had poor skills in identifying a clear focus or their perception of the rhetorical problem may be partially off-topic, especially for low-proficiency writers (Guo & Wang, 2004; Liang, 2006). Some students may drift away from the topic by addressing only part of the writing prompt. Furthermore, when approaching the topic, students' thesis statement or main argument was sometimes unclear/too-general or non-existent (Lee & Deakin, 2016; Yu, 2012). Current studies suggest three main factors causing Chinese students' straying away from the topic: (1) planning strategy in identifying the rhetorical problem, (2) lack of individual voice awareness (Ouyang & Tang, 2006), and (3) overemphasis on grammar over content (Ye, 2017). For

instance, Ye (2017) related this lack of focus and relevance to students' over emphasis on grammar accuracy in writing.

Concerning global coherence, the common problems were weak introduction or conclusion (Chong et al., 2014), and paragraph too long or too short (Hinkle, 2011). For instance, Chong et al. (2014) generated a list of 26 error types from IELTS writing scripts and found that students' paragraphs being too long or too short ranked among the top five errors. Students might include too much information in one paragraph, which makes the paragraph too long, or might fail to provide appropriate supporting details, which considerably shortens a paragraph. Hinkel (2011) presented a systematic overview of research conducted on L2 writing from 1950 and analysed the problems that L2 writers are likely to face. Four typical problems with global coherence were synthesized in her study. First, L2 writers tend to use discourse moves and their contents inconsistently, primarily due to the negative transfer of discourse-structuring conventions across various cultures. Second, they might over- or under-estimate the amount of background knowledge of readers and the need for textual clarity, explicitness and specificity. Thirdly, they might neglect to account for counter-arguments.

Fourthly, they might delay or omit thesis/main point statements as well as omit or dramatically shorten conclusions/closings (such as the use of one-sentence closings). As a result, she stated that 'L2 writers need intensive and extensive instructions in practically all aspects of constructing discourse and reasonably fluent and accurate text' (Hinkle, 2011, p. 535).

Regarding local coherence, typical problems identified were coherence breaks or unrelated idea progression in the connection of ideas (Liang, 2006; Green et al., 2000; Hong & Xu, 2016; Yu, 2012) and underuse of counter-arguments (Qin & Karabacak, 2010). For instance, Yu (2012) found that students sometimes write a topic sentence at the beginning but fail to develop it. They might shift the topic without any purpose, which can affect the information flow and hinder the logical connection. Hong and Xu (2016) observed that although students can demonstrate coherent development of ideas through adjacent sentences in their writing, the coherence in-between often lasts no longer than three sentences. Green et al. (2000) indicated that Chinese writers tend to place certain topic-fronting devices and logical

connectors in the sentence-initial position to introduce new information. This theme position negatively affects the information structure and both local and global text coherence.

Regarding the use of logical connectives, the common problems were overuse, underuse, or restricted range of certain connectives (Bolton et al, 2002; Hu & Li, 2015). For instance, among all kinds of conjunctions, it was found that resultative (e.g., *so, thus, therefore*) and listing connectors (e.g., *for example, firstly, also*) were frequently found to be overused by Chinese learners (Hu & Li, 2015). Bolton et al (2002) identified Hong Kong undergraduates' top 10 most overused connectors, with their differences from the academic norm: *so, and, also, thus, but, therefore, moreover, then, on the other hand,* and *in fact.* Some academic connectors are not used at all by the students (*on the whole, on the one hand, in contrast, in sum, in the event, in total, or, still*).

Regarding reader–writer interaction, hedging devices have been receiving increasing attention in academic writing as a prominent means to show writers' stance and convey commitment and degree of engagement (Hinkel, 2005; Knoch, 2009; Lee & Deakin, 2016).

Numerous studies have discussed the use of hedges in comparison with the features of boosters. Many studies have indicated that Chinese L2 learners tend to overuse boosters and underuse hedges in academic writing (Hinkel, 2005; Hyland, 2012; Hyland & Milton, 1997). Knoch (2009) found that in undergraduate student writing of relatively short essays, there are very few uses of attitude markers, commentaries and markers of writer identity. Instead, with an increase in their writing proficiency, students tend to use more hedges and fewer boosters.

2.3 Research design and methods

The overall question this study aimed to address is: What specific problems do students face in terms of DC in English academic writing? The question is further divided into three sub-questions to guide the study.

RQ1a. What are students' overall discourse strengths and weaknesses in their essays?

RQ1b. What are students' specific strengths and weaknesses within each discourse category?

RQ1c. Are there any similarities or differences among students with high, medium, and low writing proficiency?

2.3.1 Writing context

This study was conducted in a China-foreign joint degree program at a university in mainland China. The university has an approximately 110-year-old history, and its international program has approximately 20 years of substantial cooperation experience with overseas universities in English as native language countries. The teaching staff in this program comprises foreign and local language teachers and foreign and local subject teachers. They share offices, meeting rooms and teaching venues. Furthermore, many textbooks and teaching materials are introduced from abroad. To meet the curriculum needs and achieve academic excellence, the students are expected to compose academic essays in English for both their language courses and disciplinary courses. In addition, the students need to satisfy the curricular expectations of both the home university and the partner university to complete their degree studies. Altogether, the abundant disciplinary writing exercises in the research site provide a meaningful venue for the present study to gain detailed insights into EFL learners' genuine academic writing situations.

2.3.2 Focal participants

The focal participants were a group of business undergraduates in their second year of study. They were of similar ages and most of them shared similar learning goals. Most students chose their major from their first year.

Stratified sampling was adopted to obtain a representative sample of business undergraduates from each of the three writing proficiency levels (High-, Medium- and Low-proficiency). The program has a total of 329 students (in nine classes) majoring in business and accounting studies. These students wrote on three topics. The subject teacher's original grades were adopted as the external criterion to categorize students into three proficiency groups: low (scoring 13≤), medium (scoring 14~15) and high (scoring ≥16) within a total score of 20. Equal numbers of students (N=12) were randomly selected from each proficiency levels. This selection resulted in 108 sample essays. The detailed sampling procedures are tabulated below in Table 2.2 below.

Table 2.2

Sampling framework and sample

Writing topic	# students		LP	MP	HP	Sample
Cultural clash in the boardroom	Class 1	37	4	20	13	36/112
	Class 2	37	5	19	13	
	Class 3	38	6	19	13	
	Sample		12	12	12	
Basketball semi-final	Class 4	39	12	22	5	36/111
	Class 5	37	4	26	7	
	Class 6	35	6	17	12	
	Sample		12	12	12	
The plagiarised assignment	Class 7	36	4	22	10	36/106
	Class 8	36	4	14	18	
	Class 9	34	13	10	11	
	Sample		12	12	12	
Total	329		58	169	102	108

This sampling process generated 108 essays on three parallel prompts written by 108 students at three writing proficiency levels.

2.3.3 Design and development of the discourse rubric

The proposed discourse rubric comprised five main constructs and ten sub-categories, which were assessed along five levels. A brief overview of the discourse rubric is shown in

Appendix A. The rubric adopted the five-point scale and adjusted the descriptors into three categories for data analysis. That is, in the five points, categories 4–5 were classified as ‘good/excellent’, 2–3 were classified as ‘basic/fair’ and ‘1’ was classified as ‘poor’. ‘Poor’ suggests that there is strong/salient sense of lack of certain discourse awareness as detected from the written script; ‘basic/fair’ indicates that the written script somehow shows awareness of the key discourse features and ‘good/excellent’ indicates that the students demonstrate clear awareness and knowledge of the key discourse features, but the language might not be concise and precise.

There are four main reasons for adopting a five-point scale. First, theoretically, the five levels symbolise a wider range of potential abilities that the students need to develop. Second, as suggested by Bachman and Palmer (2016), using more scaling points than there are decisions can enhance inter-rater reliability because individual raters might not necessarily agree on exact scales. Third, for classroom use, classroom teachers also preferred the five-point scales in that the scores generated can be easily integrated into the overall course assessment in terms of percentages.

The descriptors at most levels included both qualitative and quantitative/numeric criteria for validity and consistency concerns. The qualitative descriptions and quantifiers for each category were developed with reference to the results obtained from previous rubric validation studies (Knoch, 2009; Zhao, 2010). Then, the quantifiers were modified based on a trial analysis of benchmark essays. Afterwards, the proposed rubrics and quantifiers in all categories were discussed with three local language/writing teachers and two subject teachers with reference to previous students' written performances.

2.3.4 Data preparation

Before applying the proposed assessment rubrics, an assessment manual was developed, which comprised the following five main parts: (1) operational definitions and elaboration of the 10 assessment constructs, (2) a set of codes for analysing the essays, (3) a set of typical benchmark essays for reference, (4) suggested coding procedures and (5) continual discussion on evaluation criteria during the coding and evaluation processes. In addition, the basic knowledge of the key concepts in the disciplinary domain, as assessed in the writing task [i.e., cultural/ethical dilemma and problem-solving techniques (typically Kepner-Tregoe

analysis and Dunker diagrams)], was also introduced for clearer interpretations of the students' writing. The manual was used as the core reference guide for assessing the 108 essays.

To ensure inter-rater consistency, the two coders first assessed the essays on the first four codes {[F1]–[F4]}, followed by categories [F5] and [F6] and then code [F7]. After finishing the first seven categories, the remaining three items, logical connectives {[F8] and [F9]} and complexity of hedges and boosters {[F10]}, were coded. Related textual features were analysed based on a finite set of collected features, and the two coders performed the research and evaluation work simultaneously for each essay. The internet browser (Firefox) was identified as an assisting technique to perform multiple searching and highlighting tasks. With automatic identification of related words, the coders noted these identified words in an Excel document, discussed their use and reached an agreement immediately after each search.

Regarding the rating process, two coders had a trial assessment with nine essays before a large-scale assessment. The essays were assessed in three sessions—items [F1]–[F4] in session 1, [F5]–[F7] in session 2 and [F8]–[F10] in session 3—with each session lasting approximately four hours. All sessions were conducted face-to-face to enable immediate discussion of questions. At the end of the trial assessment, Cohen's k was run to calculate the inter-rater reliability. When the two coders achieved a kappa value exceeding .6, which is considered sufficient (McHugh, 2012), they moved on to assessing the remaining essays. For approximately every hour, the coders were reminded to take a break to avoid potential exhaustion. Based on the results of the assessment of the 108 essays, the preliminary rubric was revised and then an expert's judgment of the new rubric was sought.

2.3.5. Data analysis

After completing the assessment of all 108 essays, Many-facet Rasch Measurement (MFRM) was run with FACET 3.71.3 to validate the inter-rater reliability (Linacre, 2013). The analysis focused on the first seven features {[F1]–[F7]}. As the results for [F8]–[F10] were primarily obtained through automatic search, they were less subjective and more reliable. According to the statistics reported in FACET, the inter-rater agreement opportunities were 756, with 506 =

66.9% exact agreements and 273.3 = 36.1% expected; therefore, the overall Kappa coefficient was only 0.48 = (506 − 273.3)/(756 − 273.3). The inter-rater correlations were 0.68–0.70.

This moderate level of Kappa value was considered acceptable, as it is often impossible to achieve an exact agreement in writing assessment using rubrics with multiple-point scales (Bachman & Palmer, 2016; Weigle, 2002). In practice, a one-point difference at the multiple-point scale can be considered as agreement in writing assessment (Zhao, 2010). Calculating the percentage agreement (ratings that were the same or within one-point difference) showed high inter-rater reliability: 90% for [F1]; 96% for [F2]; 94% for [F3] and 98% for [F4]–[F7]. Hence, the degree of the present inter-rater reliability is acceptable.

The scores were then subjected to a one-way ANOVA test to compare students at three proficiency levels in terms of the discourse features in their essays. The statistical assumptions of ANOVA were examined through boxplot for outliers, the Shapiro–Wilk test for normality, Pearson correlation coefficients between the dependent variables for detecting

multicollinearity, scatterplot matrix for linearity and Levene's test for homogeneity. A total of 21 outliers were identified for 11 variables at the three levels. Given the small sample size, the outliers were retained. Most of the 11 variables were not normally distributed, as assessed by the Shapiro–Wilk test ($p < .05$). As a result, a non-parameter test and a post-hoc test (Tukey's honest significant difference) were conducted. The following section reports the findings of the analysis.

2.4 Findings

This section first reports the overall discourse strengths and weaknesses based on the mean values of the five-point scaling and head distribution analysis. Then, the students' specific strengths and weaknesses within each discourse category are illustrated with excerpts from the evaluated essays. Afterwards, the similarities or differences at the three writing proficiency levels are reported.

2.4.1 Overall discourse strengths and weaknesses

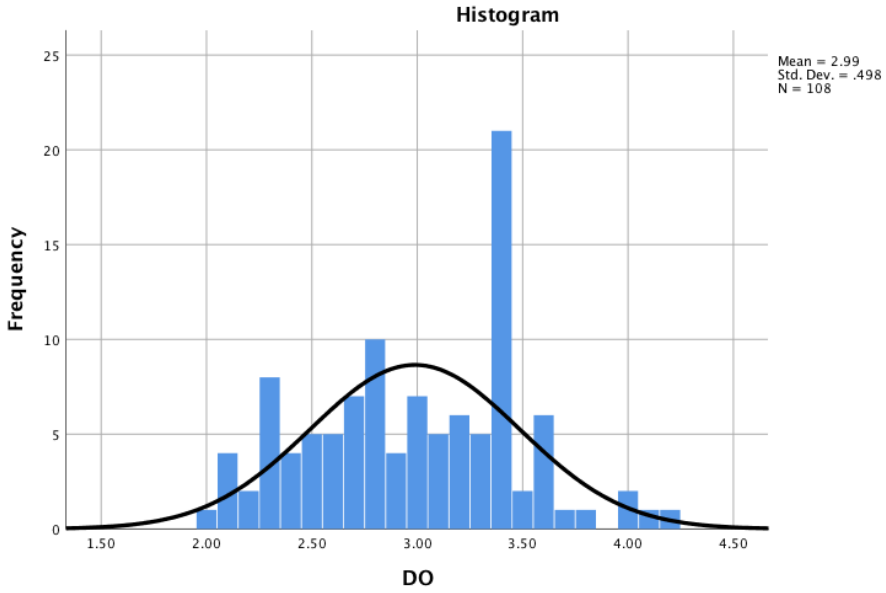
RQ1a. *What are students' overall discourse strengths and weaknesses as shown in their written essays?*

Students' discourse features were evaluated with a five-point scale. To perceive the overall situation, their writing performance was analysed in terms of overall distribution pattern (Graph 2.1 below), average mean values for each category (Table 2.3 below) and head distribution features (Graph 2.2 below).

The discourse overall (DO) was computed by averaging the values of the 10 sub-categories. The mean value distribution for DO is shown in Graph 2.1. The distribution of students' DO scores approximate a normal distribution (mean = 2.99, standard deviation =.498, skewness = .009), with the exception of the spike at 3.40 (for 19% of the participants). The minimum value was 2.00 and the maximum value was 4.20. Further check of these 19% participants found that their scores at [F8] and [F9] made salient contribution to the overall discourse score of 3.40. Overall, they demonstrated fair discourse features at the medium level. They had some basic discourse knowledge and skills but still larger space for further improvement to achieve an excellent level of 5.00.

Graph 2.1

Histogram of students' DO performance



According to Table 2.3 below, of the 10 sub-categories, the top three features were [F9] connective accuracy, [F8] connective complexity and [F1] topic/focus, and the bottom three features were [F3] controlling idea, [F6] conclusion and [F2] thesis statement. That is, the students showed certain strengths in logical connectives but were relatively weak in topic building and global coherence, particularly in terms of thesis statement and conclusion.

Table 2.3

Descriptive statistics of DO performance

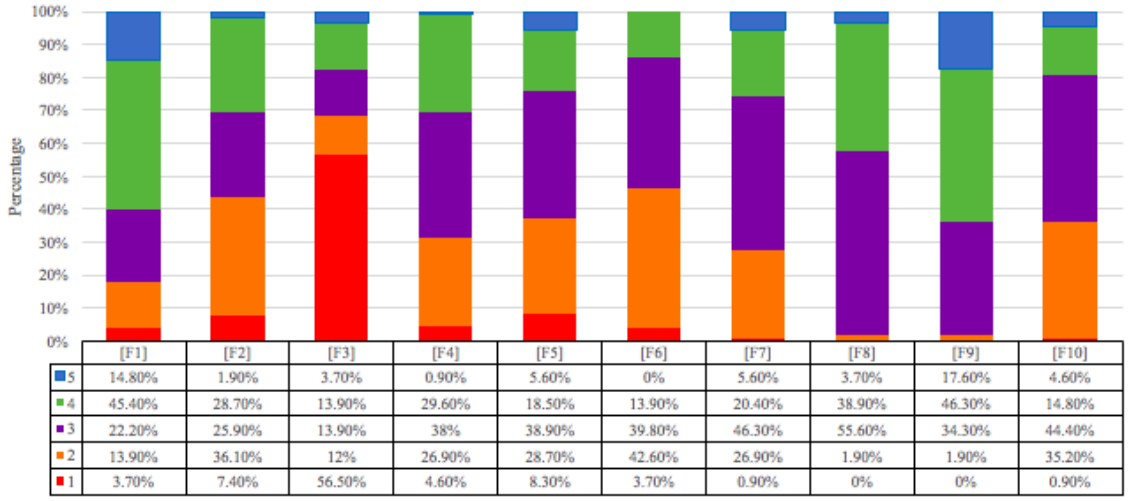
Category	Minimum	Minimum	Mean	Std. Deviation	Skewness	Kurtosis
[F1]	1.00	5.00	3.54	1.03	-0.60	-0.18
[F2]	1.00	5.00	2.81	1.00	0.04	-0.98
[F3]	1.00	5.00	1.96	1.27	0.94	-0.53
[F4]	1.00	5.00	2.95	0.90	-0.23	-0.68
[F5]	1.00	5.00	2.84	1.01	0.16	-0.34
[F6]	1.00	4.00	2.64	0.77	0.22	-0.55
[F7]	1.00	5.00	3.03	0.86	0.40	-0.18
[F8]	2.00	5.00	3.44	0.60	0.48	-0.17
[F9]	2.00	5.00	3.80	0.75	0.08	-0.67
[F10]	1.00	5.00	2.87	0.84	0.63	0.09
Discourse Overall	2.00	4.20	2.99	0.50	0.01	-0.66

Furthermore, head distribution of the 10 discourse features was computed to achieve a clearer picture of the students' relative strengths and weaknesses. Graph 2.2 displays the students' scaling values in coloured columns: red for students scaling at 1 on the 5-point scale, orange for those scaling at 2, purple for those at 3, green for 4 and blue for 5. As shown in the graph,

more than half the students were rated 4 or 5 on [F1], [F8] and [F9], whereas more than 80% of the students were rated 3 or below for [F3], [F6] and [F10]. For the other four sub-categories {[F2], [F4], [F5] and [F7]}, approximately 70% of the students were rated at 3 or below. The head distribution results in Graph 2.2 below correspond with the descriptive statistics shown in Graph 2.1 and Table 2.3 above. That is, of the five main discourse categories, the students had relative strengths in logical connectives but were weaker in topic building, global coherence and complexity of hedges and boosters.

Graph 2.2

Head distribution of the 10 discourse features on a 5-point scale



	[F1]	[F2]	[F3]	[F4]	[F5]	[F6]	[F7]	[F8]	[F9]	[F10]
5	14.80%	1.90%	3.70%	0.90%	5.60%	0%	5.60%	3.70%	17.60%	4.60%
4	45.40%	28.70%	13.90%	29.60%	18.50%	13.90%	20.40%	38.90%	46.30%	14.80%
3	22.20%	25.90%	13.90%	38%	38.90%	39.80%	46.30%	55.60%	34.30%	44.40%
2	13.90%	36.10%	12%	26.90%	28.70%	42.60%	26.90%	1.90%	1.90%	35.20%
1	3.70%	7.40%	56.50%	4.60%	8.30%	3.70%	0.90%	0%	0%	0.90%

2.4.2 Specific strengths and weaknesses within the 10 categories

RQ1b. *What are students' specific strengths and weaknesses within each discourse category*?

Students' relative strengths lay in [F1] topic/focus, [F8] connective complexity and [F9] connective accuracy. In particular, approximately 60% of the students achieved a scale of 4 or 5 on [F1]. That is, in these student essays, a thesis was established, often in concise and consistent terms (as perceived from the frequencies of the keywords, or their antonyms, synonyms, hypernyms or hyponyms) at the beginning of writing, indicating the specific focus of the essay. For instance, in the introduction of NO. 3 Essay, the keyword 'problem' and its synonyms 'dilemma, difficulty, question and conflict' appeared more than four times and thus presented a clear and consistent focus. However, this group of words was used somewhat redundantly, thus achieving an overall scale of 4, rather than 5.

> *'In the face of business cultural dilemma and ethical dilemma, it should find out the underlying problem firstly. A number of cultural aspects influence the way how to solve challenges, problems, and conflicts (Heggertveit, 2012). An ethical dilemma is a conflict between alternatives where no matter what a person does, some ethical principle will be compromised (Hedge, 2019). When Liu confronted this situation, the problem may be is how to solve the conflict between bribery and incentives, whether the*

company should bribe Wang, or what things are bribery. Through the KT situation analysis, it can be seen that timing, trend and impact are high in the first problem. Therefore, it can be included that the most important problem is how to solve the conflict between bribery and incentives.' (Excerpt from NO. 3)

The second strength was related to the use of logical connectives [F8] (with 42.6% students were scaled to 4 or 5) and [F9] (with 63.9% students were scaled to 4 or 5). Logical connectives were assessed with a prepared list of connectives in four main categories: additive, adversative, causal and temporal. The statistics of the students' use of connectives showed that all students could use connectives of at least three categories, and these connectives were grammatically correct and semantically appropriate. However, only 3.7% of the students were scaled to 4 rather than 5 on [F8]. Further examination of the connective statistics showed that among the four connective categories, the students often preferred the additive and causal categories, with less use of adversative and temporal categories. Among these four categories, the most frequently used words included 'and', 'also', 'too', 'because',

‘so’, ‘if’, ‘therefore’, ‘firstly...’ ‘then’ and ‘in conclusion’. In other words, connective complexity requires further improvement.

The weakest features were [F3], [F6] and [F10], of which [F3] was the weakest. More than half the students (56.5%) neither had a sentence nor any complementary information to present a brief navigation of the whole essay in the introduction. This indicates a lack of common awareness regarding the significance of controlling ideas in the introduction.

The second weakness was related to [F6], with no one scoring 5 but 42.6% scoring 2. Based on the task requirements and conventional features in the conclusion, four components were considered essential: (1) writer’s position, (2) presence of justification/explanation/analysis, (3) quality of justification/explanation/analysis and (4) further suggestions. The essays which were scaled to 1 on [F6] tended to end abruptly, with no signals to give readers a sense of closure. Among the essays which were scaled to 2 on [F6], some presented too short conclusions, for instance, with only one sentence ‘*To sum up, the last choice seems to be the best answer to the problem*’ (Excerpt NO. 7). Some essays had conclusions with very limited

information, with occasionally unnatural insertion of citations; for example, '*In conclusion, 'let they join in the game with more practice' is the best solution, because it 'suits the all must', the interest of the team, rule, and fairness. According to Wang (2000), 'create a fair environment is very important for a team, which will build a creative group'* (Excerpt NO. 14).

The third weakness was related to [F10]. Approximately 36% of the students were scaled to 2 or below and 44.4% were scaled to 3. Typical expressions of hedges and boosters were coded based on a prepared list of expressions. If neither hedges nor boosters were used, the essay was scaled to 1. If more boosters than hedges were used, or if there was severe disproportion, the essay was scaled to 2. If the use of hedges or boosters included mostly common modal verbs such as 'may(be)', 'must', 'should' or 'will', the essay was scaled to 3. Statistics of the coding results indicated that most students (80%) did not appropriately use the hedges and boosters. Most of them either tended to use more boosters than hedges or only used very common modal verbs. Thus, there is further need for sophistication and variation.

The remaining four sub-categories (i.e. [F2], [F4], [F5] and [F7]), although not much saliently weak, were also not very good. For instance, regarding [F7], many cases of coherence break and unrelated idea progression shared the feature of including a citation or reference. This suggests that the students wrote to fulfil the citation requirement, or they were not clear why, what and how to integrate the source information.

> *'To control emotion with reason requires the leaders to have a correct social cognition and value or intention.' (Shujun, Z., 2013). What should coach Jeff do as the leader of this basketball team?*
>
> *There are two plans here:*
>
> *Plan A: Punish immediately. It can ban two athletes and maintain the principle of fairness, and also justice.*
>
> *Plan B: Punish them after the semi-finals. The result of competition is more important.*
>
> (Excerpt from Essay NO. 32)

In addition, many students' in-text citations seemed to have been directly translated from Chinese to English, probably through online translation. Hence, there was much concern about the linguistic quality and content clarity of the citation information. This also posed a considerable threat to local coherence.

2.4.3 Similarities or differences at three levels of English writing proficiency (RQ3)

Table 2.4 presents the descriptive statistics; Table 2.5 presents results of the non-parameter tests, and Table 2.6 presents the results of multiple comparison analysis. As shown in Table 2.4, HP students, on average, displayed better overall discourse features than MP and LP students (DO means = 3.194, 2.878 and 2.894 respectively, $\chi 2(2) = 9.42$, $p = 0.01$). The differences between HP and MP (p =.025) and HP and LP (p =.017) are statistically significant, that between the MP and LP is not.

Similarly, HP students significantly outperformed MP and LP students in terms of conclusion writing (F6: χ 2(2) = 6.88, p = 0.03) and connective complexity (F8: $\chi 2(2) = 6.29$, p = 0.04). However, only the differences between the HP and LP groups are significant; the differences between HP and MP are not. Finally, no significant differences were found between MP and LP students in any of the 10 features.

Table 2.4

Descriptive statistics of the similarities or differences among the three proficiency levels

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
						Lower Bound	Upper Bound		
F1	LP	36	3.36	1.150	.192	2.97	3.75	1	5
	MP	36	3.39	.964	.161	3.06	3.72	1	5
	HP	36	3.86	.899	.150	3.56	4.17	1	5
F2	LP	36	2.86	1.073	.179	2.50	3.22	1	5
	MP	36	2.75	.967	.161	2.42	3.08	1	5
	HP	36	2.83	.971	.162	2.50	3.16	1	4
F3	LP	36	1.81	1.215	.202	1.39	2.22	1	5
	MP	36	1.75	1.156	.193	1.36	2.14	1	4
	HP	36	2.33	1.373	.229	1.87	2.80	1	5
F4	LP	36	2.86	.931	.155	2.55	3.18	1	4
	MP	36	2.94	.826	.138	2.66	3.22	1	4
	HP	36	3.06	.924	.154	2.74	3.37	1	5
F5	LP	36	2.75	.937	.156	2.43	3.07	1	5
	MP	36	2.64	1.073	.179	2.28	3.00	1	4
	HP	36	3.14	.961	.160	2.81	3.46	2	5
F6	LP	36	2.42	.692	.115	2.18	2.65	1	4
	MP	36	2.61	.838	.140	2.33	2.89	1	4
	HP	36	2.89	.708	.118	2.65	3.13	2	4
F7	LP	36	3.00	.756	.126	2.74	3.26	2	5
	MP	36	2.78	.797	.133	2.51	3.05	1	4
	HP	36	3.31	.951	.158	2.98	3.63	2	5
F8	LP	36	3.25	.604	.101	3.05	3.45	2	5
	MP	36	3.50	.609	.102	3.29	3.71	3	5
	HP	36	3.58	.554	.092	3.40	3.77	3	5

F9	LP	36	3.69	.749	.125	3.44	3.95	2	5
	MP	36	3.75	.649	.108	3.53	3.97	3	5
	HP	36	3.94	.826	.138	3.66	4.22	3	5
F10	LP	36	2.94	.984	.164	2.61	3.28	2	5
	MP	36	2.67	.756	.126	2.41	2.92	2	5
	HP	36	3.00	.756	.126	2.74	3.26	1	4
D-	LP	36	2.894	.544	.091	2.71	3.08	2.00	4.20
O	MP	36	2.878	.460	.077	2.72	3.03	2.10	4.10
	HP	36	3.194	.431	.072	3.05	3.34	2.20	4.00
	Total	108	2.989	.498	.048	2.89	3.08	2.00	4.20

Table 2.5

Results of the non-parameter test on the similarities or differences among the three proficiency levels

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	DO
Kruskal–Wallis H	5.296	.338	5.027	.835	3.674	6.877	5.424	6.288	1.529	4.393	9.420
Df	2	2	2	2	2	2	2	2	2	2	2
Asymp. Sig.	.071	.845	.081	.659	.159	**.032**	.066	**.043**	.466	.111	**.009**

Table 2.6

Similarities or differences among three proficiency levels

Dependent Variable	(I) Level	(J) Level	Mean Difference (I-J)	SE	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
F1	Low	Mid	-.027	.238	.993	-.594	.538
		High	-.500	.238	.095	-1.066	.066
	Mid	Low	.028	.238	.993	-.538	.594
		High	-.472	.238	.121	-1.038	.094
	High	Low	.500	.238	.095	-.066	1.066
		Mid	.472	.238	.121	-.094	1.038
F2	Low	Mid	.111	.237	.886	-.452	.674
		High	.028	.237	.992	-.535	.591
	Mid	Low	-.111	.237	.886	-.674	.452
		High	-.083	.237	.934	-.647	.480
	High	Low	-.028	.237	.992	-.591	.535
		Mid	.083	.237	.934	-.480	.647
F3	Low	Mid	.056	.294	.981	-.646	.757
		High	-.528	.295	.178	-1.229	.173
	Mid	Low	-.056	.295	.981	-.757	.646
		High	-.583	.295	.123	-1.285	.118
	High	Low	.528	.295	.178	-.173	1.229
		Mid	.583	.295	.123	-.118	1.285
F4	Low	Mid	-.083	.211	.918	-.585	.418
		High	-.194	.211	.628	-.696	.307
	Mid	Low	.083	.211	.918	-.418	.585

		High	-.111	.211	.858	-.613	.390
	High	Low	.194	.211	.628	-.307	.696
		Mid	.111	.211	.858	-.390	.613
F5	Low	Mid	.111	.234	.883	-.445	.667
		High	-.389	.234	.224	-.945	.167
	Mid	Low	-.111	.234	.883	-.667	.445
		High	-.500	.234	.087	-1.056	.056
	High	Low	.389	.234	.224	-.167	.945
		Mid	.500	.234	.087	-.056	1.056
F6	Low	Mid	-.194	.176	.515	-.614	.225
		High	-.472*	.176	**.023**	-.892	-.053
	Mid	Low	.194	.176	.515	-.225	.614
		High	-.278	.176	.261	-.697	.142
	High	Low	.472*	.176	**.023**	.053	.892
		Mid	.278	.176	.261	-.142	.697
F7	Low	Mid	.222	.198	.501	-.248	.692
		High	-.306	.198	.274	-.776	.164
	Mid	Low	-.222	.198	.501	-.692	.248
		High	-.528*	.198	**.024**	-.998	-.058
	High	Low	.306	.198	.274	-.164	.776
		Mid	.528*	.198	**.024**	.058	.998
F8	Low	Mid	-.250	.139	.175	-.580	.080
		High	-.333*	.139	**.047**	-.664	-.003
	Mid	Low	.250	.139	.175	-.080	.580
		High	-.083	.139	.821	-.414	.247
	High	Low	.333*	.139	**.047**	.003	.664
		Mid	.083	.139	.821	-.247	.414

F9	Low	Mid	-.056	.176	.946	-.473	.362
		High	-.250	.176	.332	-.668	.168
	Mid	Low	.056	.176	.946	-.362	.473
		High	-.194	.176	.512	-.612	.223
	High	Low	.250	.176	.332	-.168	.668
		Mid	.194	.176	.512	-.223	.612
F10	Low	Mid	.278	.198	.342	-.192	.748
		High	-.056	.198	.957	-.526	.415
	Mid	Low	-.278	.198	.342	-.748	.192
		High	-.333	.198	.215	-.803	.137
	High	Low	.056	.198	.957	-.415	.526
		Mid	.333	.198	.215	-.137	.803
DO	Low	Mid	.017	.113	.988	-.253	.286
		High	-.300*	.113	**.025**	-.569	-.030
	Mid	Low	-.017	.113	.988	-.286	.253
		High	-.317*	.113	**.017**	-.586	-.047
	High	Low	.300*	.113	**.025**	.031	.569
		Mid	.317*	.113	**.017**	.047	.586

Notes: * The mean difference is significant at the .05 level.

2.5 Discussion

This section first expands on the effect of construct specificity on DA, then compare the findings on students' specific problems and difficulties regarding DC with previous studies. Implications for EFL/L2 wring instruction in university are also discussed.

Construct specificity, or the grain size of constructs, is a unique problem to be resolved in DA. Both Jang (2009) and Kim (2019), for instance, considered identification of specific constructs to be the most important procedure in implementing DA. A fine-grained rubric can contribute to the interpretation of specific constructs, and enable the provision of diagnostic feedback and retention of effective student engagement. Fox et al. (2016) tackled the problem via emphasising a disciplinary (rather than generic) assessment approach as they considered disciplinarity as critical if interventions were to provide an opportunity for socialisation into disciplinary practices. While acknowledging the critical role of disciplinarity in specifying the exact criteria for assessment, it is equally important to specify the extent of grain size within specific constructs. The present study decomposed the crucial yet often elusive concept DC into five components and 10 features and assessed these features on a five-point scale. The study found this 10×5 rubric could differentiate writing at two proficiency levels and generate a nuanced diagnostic profile of discourse-level qualities in the students' academic writing. Thus, this study enriches the current understanding of construct specificity in DA research.

The study results indicated that the students, on average, exhibited a basic level of performance in discourse-level features. This supported Hinkel's (2011) statement that L2/EFL learners may need instruction in almost all aspects of discourse construction. The genre-based approach to writing can be considered a useful pedagogy in improving learner/writer DC in writing, because the core concern of this approach is about the linguistic patterns, rhetorical strategies, and shared social conventions, which resonate with our conceptualization of DC. In existing studies on EFL/L2 writing, research on EFL writing approaches typically evolve around the product, process and post-process approaches, or content-based instruction and content and language-integrated learning (Ferris & Hedgcock, 2014; You, 2004). There are relatively few published studies that specifically focus on classroom practice of genre-related approaches to writing, typically through the exploration of the genre-based approach (Tardy, 2006) and process-genre approach (Badger & White, 2000; Huang & Zhang, 2020; Racelis & Matsuda, 2013). The genre-based approach is mostly investigated in the ESP context, but remains under-researched in EFL/L2 writing practice. Hence, the development of proper language learning objectives in EFL/L2 writing courses

and the subsequent arrangement of instructional sequences form the first steps in facilitating EFL/L2 learners' development of DC in English academic writing.

One finding of this study is that these EFL learners' problems with discourse features are related to global coherence. In particular, controlling idea and conclusion writing are the weakest aspects. Typical problems with the students' conclusion writing included an abrupt ending of the essay with no explicit signal, conclusion being too short and an abrupt insertion of new information. A similar finding was reported in Chong et al. (2014), who found the top two errors at the discourse level were poor conclusion and poor introduction. Existing studies suggest two main factors contributing to global coherence problems: mother tongue influence and developmental factors (Hinkel, 2011). For instance, the Chinese culture emphasises parataxis more than hypotaxis, and Chinese people tend to be implicit in expressing themselves. As the perception of coherence depends on the writer, the reader and the text, Thompson (1986) proposed the practice of audience analysis to help the students accommodate shared cultural, professional and linguistic knowledge for more readable and coherent writing. This writer–audience approach was also recommended by Lee (2002) and

Tardy and Swales (2009) as particularly helpful in understanding rhetorical patterns in texts extracted from different national and linguistic cultures. In addition, outline planning and revision of global structure directly influence the demonstration of global coherence. Hence, planning and revision strategies are also required to facilitate student development of global coherence.

Besides global coherence, another typical problem identified in this study is with reader–writer interaction, particularly the salient features of the use of hedges and boosters. In this study, more than one third of the participants were found to use more boosters than hedges, or there was a severe disproportion of their use. In addition, there was a limited range of use, with some modal verbs as the most commonly used hedges and boosters. Similar observations of EFL/L2 students' overuse of boosters and underuse of hedges were reported in previous studies (Hinkel, 2005; Hyland & Milton, 1997). For explicit teaching and learning of these discourse features at university study, Hyland and Milton (1997) suggested teaching strategies on differentiation between the observed facts and interpretation and on both variety of expressions and potential ranges in conveying different meaning. Hinkel

(2005) emphasised the necessity of instruction on hedges in effective L2 writing. However, the present study found no significant differences between HP and MP/LP student writings in terms of the use of hedges and boosters. This suggests that at an early stage of learning in academic writing in the EFL/L2 context, the features and functions of hedges and boosters are not highly valued, at least by subject teachers in disciplinary writing practice. This is also partly in line with Stapleton's (2002) proposal of reasoning and argumentation preceding over reader-writer interaction features. In light of the above, it seems that global coherence should be the primary concern and reader-writer interaction as secondary in classroom instruction in EFL/L2 writing course.

2.6 Conclusion

This study has developed a set of discourse assessment rubrics to diagnose novice EFL writers' problems and difficulties in writing essays. The results showed that the students, on average, exhibited a basic level of discourse features. Their relative strengths were on *topic/focus*, *connective complexity* and *connective accuracy* and relative weaknesses were on *thesis statement*, *controlling idea* and *conclusion*. They also had some problems with *reader orientation*, *body paragraphs*, *local coherence* and *complexity of hedges and boosters*. Of the

five main discourse categories, the students were stronger in *logical connectives* but weaker in *topic building*, *global coherence* and *complexity of hedges and boosters*. Their performance on *local coherence* was in between. Furthermore, HP students performed significantly better than MP and LP students in terms of *conclusion writing*, *connective complexity* and *overall discourse feature*.

The results showed that the present discourse rubric is a useful instrument to identify novice L2 writers' specific strengths and weaknesses at the discourse level. Although the rubrics should be trialled on writing samples of other academic genres, especially longer pieces of writing that demand a stronger command of DC. The present textual analysis focused on one type of relatively short academic writing by the students from one discipline. Future studies should examine and compare essays written in a different genre for a different discipline to have a more comprehensive view of Chinese EFL undergraduate writers' development of DC in English academic writing.

This chapter only examines students writing problems as shown in their written scripts. To diagnose the specific strengths and weaknesses for effective treatment, it is necessary to investigate the composing process and identify pertinent factors that may contribute to particular textual problems. The following chapter, i.e., Chapter 3, conducted a comparative case study in Hong Kong and in mainland China to have deeper and richer understanding of Chinese EFL learners' English academic writing experience.

Chapter 3 Study Two — A Comparative Case Study

Chapter Abstract

This comparative case study reports a three-tiered diagnostic assessment of two groups of EFL student writers' discourse competence in terms of textual features, composing strategies, and knowledge about academic writing in university study in Hong Kong and mainland China. Data were collected over two years mainly through textual analysis, semi-structured interview, and writer log. This study found that the participants demonstrated a basic level of discourse features. Their relative strengths were on *topic/focus*, *connective complexity* and *connective accuracy*; their weaknesses were on *controlling idea*, *conclusion*, local coherence and *reader–writer interaction*. *Global coherence* was particularly a salient discourse problem for mainland participants. Two factors had an immediate influence on the development of global coherence: time management skills on planning and revision and genre knowledge. In addition, L2 proficiency was found to have both interrelated and independent relationships with the development of DC in university study. Two typical strategies adopted by the LP and MP students in the mainland context were *whole-page camera translation in English academic reading* and *question-answer or blank filling for planning/formulation*. This comparative analysis revealed some discrepancies/gaps in the current preparation program and shed light on better preparation support.

3.1 Introduction

Writing poses a constant challenge for students from primary to tertiary education. Writing in a second or foreign language is even more challenging. To understand university students' writing problems and difficulties, many existing studies have focused on the analysis of written scripts (Bruce & Hamp-Lyons, 2015; Bruce, 2016; Toraskar & Lee, 2016) or that of

student needs through surveys or interviews (Evans & Morrison, 2011; McNamara et al., 2018). However, there is a dearth of attempts focusing on discourse competence (DC) and still less research tracking the actual composing process to identify possible causes and factors for the textual problems observed. Relating students' problems with potential causes, however, is necessary for planning targeted teaching and learning activities (Alderson, 2005; Harding et al., 2015).

This chapter reports a comparative multi-case study of 15 Chinese undergraduate students, focusing on their English academic writing experience across the curriculum in their first two years of study at a university in Hong Kong and at a mainland university. The case studies in Hong Kong were our exploration of the difficulties that students, once on board of an overseas EMI program, would have, while the cases in the mainland focused on the difficulties that students had in a program that prepared them to go abroad for EMI education. This comparative analysis could reveal the discrepancies/gaps in the current preparation program and shed light on better preparation support. Specifically, three tiers of diagnosis were conducted: the first tier focused on their written scripts, the second tier focused on the processing strategies and the third tier focused on their academic writing knowledge status. Based on the analysis of writing samples, writer logs and semi-structured interviews, this study identified students' discourse-level weaknesses and investigated potential influencing factors in relation to their writing strategies and academic writing knowledge status.

3.2 Literature review

This section first reviews the existing studies on student writers' composing processes in EFL academic writing. Next, empirical studies on EFL students' current academic writing knowledge are synthesized. Then, a complex dynamic perspective on second language learning was reviewed to identify the crucial learning-relevant context factors. The review

aims to achieve a systemic understanding of related discourse problems and pertinent influencing factors.

3.2.1 Studies on EFL learners' writing processes

L2/EFL writers are considered to share composing processes and textual features similar to those shared by L1 writers in many important ways (Silva et al., 2001). Both L1 and L2/EFL lines of research have found composing a recursive process, which normally comprises planning, formulating and revising (Flower & Hayes, 1981; Roca de Larios et al., 2008; Stapleton, 2010). While accounting for its recursiveness and simultaneity features, each of the three phases can be further classified into sub-entities for a more tangible description and explanation of the composing process.

Flower and Hayes' (1981) study is often considered seminal for a detailed study of the thinking processes in writing. They adopted think-aloud protocols to examine native English speakers' writing process and proposed a cognitive model of planning, translating and reviewing in composition writing. These composing processes exhibit a hierarchical structure and can occur at any time; each process component can be embedded within another process.

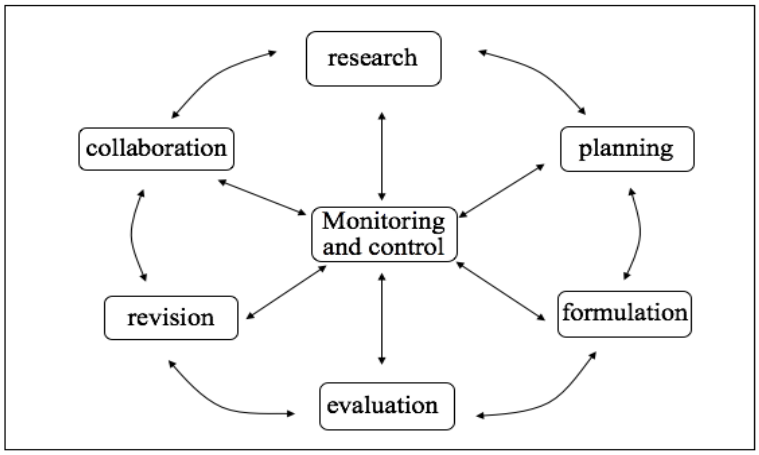
Using a similar think-aloud method but focusing on English as L2 Spanish writers' timed argumentative writing, Roca de Larios et al. (2008) conceptualised the composing process in terms of seven activities (i.e., reading the prompt, task conceptualisation, planning, formulation, evaluation, revision and meta-comments). They found that *formulation* takes up the largest portion of composition time, and writers' L2 proficiency significantly influences the time allotment across different activities; moreover, high-proficiency writers show more balanced time distribution than low-proficiency writers.

Using in-depth logs, a questionnaire and interviews, Stapleton (2010) explored the composing process of an L2 learner writing a 4000-word English essay. They proposed two additional activities in composing an academic essay: *research* (sourcing, reading and/or copying information pertaining to the composing task at hand) and *collaboration* (consulting others). Then, they analysed students' composing behaviours in these six main categories (*research*, *planning*, *formulation*, *revision*, *evaluation* and *collaboration*). They found that students allocated considerable time to *research* and much less to *formulation*, *revision* and *evaluation*. Unlike the present study, which focused on first-year undergraduate students, Stapleton (2010) focused on master's students, who have relatively higher language proficiency and a certain amount of research experience, as the research was conducted at a time when the participants had already completed their master's study. While their composing process of writing an untimed English academic essay might be similar, a close examination of L2 first-year undergraduate students' writing process might reveal more specific difficulties they encountered upon entering academia and be important to L2 academic writing instructors.

Altogether, many existing studies have integrated both L1 composition studies and L2 studies, timed as well as extended writing, by investigating individual cases. With the primary aim of understanding the process in academic writing, the present study adopts Stapleton's (2010) model as the analytical framework to probe into EFL learners' processing strategies. The typical cognitive writing activities and their interactions in the composing process are summarised in Figure 3.1 below. In this figure, the double-ended arrows indicate the interaction and recursivity of the processing strategies in the cyclical composing process.

Figure 3.1

L2 writing processing model adopted in the present study



3.2.2 EFL learners' academic writing knowledge status

Academic writing knowledge refers to the range of knowledge and skills that students must acquire to communicate competently in academic writing. Due to different perspectives on understanding the purposes and contexts of academic language, the conceptualisation of academic writing knowledge varies.

A relatively comprehensive theoretical conceptualisation of academic writing knowledge is Snow and Uccelli's (2009) four-dimensional model, which comprises the dimensions of *linguistic skills, genre mastery, reasoning/argumentative strategies* and *disciplinary knowledge.* The model is conceived to address concerns about designing instructions for academic and discipline-specific language learning in secondary and tertiary education settings. By adopting this four-dimensional model, Zhao and Lyu (2019) examined 177 Chinese undergraduate English major students' metacognitive knowledge status. They found that these students conceived academic writing as mainly comprising two general dimensions: *content knowledge* and *language skills*. These students had imprecise and vague understandings of academic writing and particularly lacked awareness of the subtle aspects of

academic genre knowledge. Similar findings were reported by Sang (2017). To understand the English writing situation in tertiary-level education in China, Sang (2017) conducted a synthesis study of 29 papers published in core and prestigious Chinese journals between 2000 and 2016. The results indicated that the emphasis is on specific linguistic knowledge rather than organisation and logical reasoning; the three main contributors were the negative influence of the curriculum and syllabus, the high-stakes standardised test and the language environment.

Within the area of L2 writing assessment, the rating scales or rubrics are considered as the de facto constructs of assessment, which convey the underlying theoretical framework supporting the assessment (Knoch, 2011). Two recent publications have reported their rubrics for assessing EFL/L2 learners in university studies: Bruce and Hamp-Lyons (2015) and Kim (2019). Bruce and Hamp-Lyons (2015) proposed an assessment rubric specially designed for academic writing in EMI universities. Their rubric assessed four main constructs, namely, task fulfilment, discourse competencies, language competencies and source integration. However, they conceptualised that there might be different degrees of overlap between task fulfilment and discourse competencies (e.g., in relation to an organisation). Kim (2019) developed and validated a list of 35 descriptors in five dimensions for assessing ESL academic writing: content fulfilment, organisational effectiveness, grammatical knowledge, vocabulary use and mechanics. Kim's five dimensions, however, were based on the analysis of high-stakes, exam-driven TOEFL essays, which might not cover the features of extended academic writing in students' disciplinary studies.

Ruan and Chen (2017) conducted a situated qualitative study to understand students' experiences of L2 disciplinary writing at an English-medium university in mainland China.

They generated four main themes in understanding L2 disciplinary writing: (1) language use in L2 disciplinary writing, (2) perceptions of rhetorical norms and disciplinary conventions, (3) valued practices in learning to write in the disciplines and (4) the role of writing in developing disciplinary competences. They found that students in their final year still considered grammar and vocabulary as the main concerns of language use in their writing. Moreover, they often depended on translation to integrate knowledge from Chinese sources and construct it in written English. Moreover, students did not have genre knowledge in relation to disciplinary conventions. Regarding writing to learn disciplinary concepts and content, students strongly believed that the manner of writing could facilitate their disciplinary competence. However, this study only probed students' writing knowledge through interviews but did not examine the genuine writing tasks and the real composing process, which could provide a more complete picture of students' mastery of academic writing knowledge.

3.2.3 A complex dynamic perspective on Second language learning

Complexity theory is a systems theory and it offers a conceptual framework to inform the way we investigate and understand the nature and the world we live in (Cilliers & Preiser, 2010; Larsen-Freeman, 2017). In specific, complexity theory involves "thinking about the social world and its intersections with the natural world as involving dynamic open systems with emergent properties that have the potential for qualitative transformation" (Byrne, 2005: 98). It is dynamic in that the complex systems may undergo periods of relative stability, but they are perpetually dynamic, having the potential to undergo radical change at any time. It is open in that the complex dynamic systems take in and expend energy, matter, or information, depending on the type of system, all the while showing the emergence of order (self-organization). It is adaptive in that the complex systems change in response to changes in its environment. That is, the adaptive systems can "learn" as a result of experience. It is

complex, rather than complicated in that its components (elements or agents) are both interconnected and spatially/temporally context dependent (Juarrero, 2000: 26).Complex dynamic systems also exhibit nonlinearity. It is nonlinear in that small change in one parameter can have huge implications downstream, which makes precise prediction unreliable or compromised. In addition, there are multiple pathways by which the system can evolve and "the same 'cause' can, in specific circumstances, produce different effects" (Urry, 2005: 4). These characteristics – "dynamic, open, adaptive, situated, interconnected, and non-linear"- provide useful theoretical lens and tools for Second language development researchers to trace and to describe emerging patterns in dynamic systems in order to explain change and growth in language and language development (Larson-Freeman, 2017).

Closely related to complexity theory is the ecological perspective on studies in (second) language learning. The term "ecology" was originated in biology studies, referring to the totality of relationships of an organism with all other organisms with which it comes into contact (Arndt & Janney, 1983; van Lier, 2004). That is, ecology studies organisms in their relations with the environment. Relating to linguistic study, an ecological perspective views language as a system of relations instead of a collection of objects and it views language learning as ways of relating more effectively with people and the world. In specific, emergence and affordance are two crucial concepts in ecology (van Lier, 2004). Affordance focuses on the relationship between an organism and the environment. Emergence refers to the reorganization of simple elements in complex systems. The complex dynamic order within an ecologic system provides affordances for active participants in the setting, and learning emerges as part of affordances being picked up and exploited for further action. In the present study, emergence and affordance will be used for interpretation of the similarities and differences.

Framing individual learning as complex dynamic systems, Dörnyei and Ryan (2015) proposed a set of four stable and systematic factors that mediate individual learning process and performance in second language development: aptitude, motivation, learning style, and learning strategies. In specific, language aptitude relates to the cognitive dimension, referring to the capacity and quality of learning. Motivation concerns the affective characteristics of the learner, referring to the direction and magnitude of learning behavior in terms of the learner's choice, intensity, and duration of learning. learning style refers to manner of learning and learning strategies engage the learner's proactiveness in selecting specific made-to-measure learning routes. In addition, they also summarized five other learner characteristics, including creativity, anxiety, willingness to communicate, self-esteem, and learner beliefs. From the complex dynamic perspective, these attributes are not static, but may show salient temporary and situational variation; they are also not monolithic but interact with each other and with the environment synchronically and diachronically (Dörnyei, 2017). Dörnyei and Ryan's (2015) dynamic and systemic frame of individual characteristics provide a relatively holistic view in understanding L2 learners' processing behaviors in English academic writing.

3.3 Research design and methods

This multi-case study addressed the following overall question (RQ2): How do EFL learners approach discourse-level features in composing processes? This question is further divided into three sub-questions to guide the study.

RQ2a. What discourse problems do these EFL undergraduate student-writers display in their English essays and what are the similarities/differences regarding their discourse features if any?

RQ2b. How do these problematic features arise in composing processes? What are the similarities/differences regarding their processing strategies if any?

RQ2c. To what extent do students' processing strategies and academic writing knowledge status affect their choices and decisions during writing? What other factors may contribute to such similarities/differences?

3.3.1 Teaching and learning context

This comparative study investigated two groups of Chinese EFL learners studying in an English-medium university in Hong Kong and in a bilingual education program in a mainland uiversity. Hong Kong has adopted biliteracy (Chinese and English) and trilingualism (English, Putonghua and Cantonese) in its education policy since its handover to China in 1997 (Li, 2017). Universities, therefore, expect their graduate students to have a high level of competence in spoken English, Putonghua and Cantonese, as well as in written Chinese and English. All eight public universities in Hong Kong adopt English as the medium of instruction for most courses. In the university where the present research was conducted, undergraduate students are required to complete approximately 30 courses in four years to obtain their first degree; that is equivalent to four to five courses each semester. Most of these courses assess students' learning achievements through written assignments in English or Chinese. Language enhancement courses focusing on English for academic purposes and general education foundation courses are offered in Year 1 and 2 studies.

The mainland case study tracked students' composing process in preparing for a 500-word essay in their disciplinary course titled *Integrated Business Challenges* (IBC). The details of the essay task is shown in Appendix D. The IBC course is delivered over 15 weeks by three lecturers in two forms: one local Chinese teacher and one foreign teacher from a partner

university for a big-class lecture and one local Chinese teacher for intensive small-class learning. In the first lecture, the course guide and key teaching, learning and assessment materials were introduced both in the online portfolio and in hardcopy. This course asked students to finish four major writing tasks, i.e., two reflection journals, a 500-word essay, and a 2000-word business report. The writing prompt, rubrics, sample essays and reading materials were provided. Then, at week 12, the subject teacher provided a two-page writing guide and had a two-hour lecture demonstrating this task with a specific case. The 500-word essay task asked students to analyse a business case by applying a series of problem-solving principles techniques. Besides reading of the provided case study materials, four to five sources were required. This 500-word essay was chosen as the present focus mainly for three reasons. First, the essay type is a typical writing task in university study across courses. Second, this essay is individual work rather than group work, and this individual work may shed more detailed light in understanding individual composing process. Third, the content aspect is related to problem-solving principles and techniques, for which the present researcher had committed half a year getting familiar with related subject knowledge, mainly by reading the textbooks and some other classical works, by attending the subject course together with students, and by individual talk with some subject teachers. Thus, the research felt more confidence in handling this 500-word essay.

Similar to the cases in Hong Kong, the writing activities occurred around the exam period in semester learning, and all students had a busy schedule preparing for various tests/exams. In particular, in the week of the 500-word essay writing, students' typical activities were an academic poster and oral presentation for an English writing course, a writing task for a 500-word journal, the nation-wide College English Test and preparation for a test in their *Mathematical Statistics* course. In addition, slightly different from the Hong Kong context,

around the exam period, all courses also require students' attendance in class as usual, and students attend about 39-h sessions sitting in class each week.

3.3.2 Focal participants

Six Chinese EFL learners majoring in business and psychology studies were recruited in an EMI university in Hong Kong. The business programme under study only recruits 1–3 students from the mainland each year. Therefore, it has a rather small body of mainland students. For this study, one Year 2 and one Year 1 mainland students were recruited from the business programme. The Year 1 mainland student was the only one from the mainland in their cohort. To enable comparison, a local Hong Kong student was also recruited. For a similar reason, three mainland students majoring in psychology were also recruited. The participants' detailed demographic information is presented in Table 3.1 below. Some details of their disciplinary writing prompts across curriculum were provided in Appendix B.

Table 3.1

The six participants' background information

Category	Participants					
	Rachel	Lydia	Chloe	Tiffany	Wei	Lucia
Gender	Female	Female	Female	Female	Female	Female
English proficiency	MP	HP	MP	MP	MP	MP
GK English	124	142	NA	130	128	126
Year of Study	Year 2	Year 1	Year 1	Year 2	Year 2	Year 2
Major	Business	Business	Business	Psychology	Psychology	Psychology

(MP: medium English proficiency; HP: high English proficiency)

On the mainland side, nine cases were selected from one teaching class based on their English proficiency level, as recommended by their language teacher and peers (three students at each of the high, medium, and low proficiency levels) and gender balance (three

male and six female). They were all Year 2 undergraduate students majoring in business studies at the time of the research. The details of the selected participants are listed in Table 3.2.

Table 3.2

The nine participants' background information

Participant	Yang	Han	Lin	Xuan	Jones	Jian	Jie	Fan	Qiang
Gender	Female	Female	Male	Female	Female	Male	Female	Female	Male
English proficiency	LP	LP	LP	MP	MP	MP	HP	HP	HP
GaoKao English scores	109	113	119	115	115	129	133	138	135

LP: Low proficiency; MP: Medium proficiency; HP: High proficiency

Prior to university study, these Chinese EFL learners had about 15 years of Chinese-medium instruction at school. Their main exposure to English writing then was a 250-word exam-oriented essay, and they had no experience in writing essays longer than that, particularly 2500-word disciplinary essays. Some details about the focal participants' English learning trajectories can be seen in Figure 3.2 below. That is, these students entered the university without any grasp of the core academic conventions used in writing research essays.

Figure 3.2

The focal participants' English learning trajectories

Age	Stage of study	Formal commitment in English	Possible English writing tasks
…			
24 23	Graduate school	…	…
22 21 20 19 18	Focal participants' undergraduate study (4~5 years)	English-medium instruction for most courses (about 15 hours/week)	Describe, report, analyze, criticize, propose, discuss (from 250 to 2500 words or longer)
17 16 15	Senior high school (3 years)	Main focus on vocabulary and grammar in reading and listening(about 8 hours/week)	Describe, discuss, or express opinions (around 150 words)
14 13 12	Junior high school (3 years)	Main focus on vocabulary and grammar in reading (about 6 hours/week)	Describe, comment, or explain in simple terms (around 100 words)
11 10 9 8 7 6	Primary school (6 years)	Main focus on pronunciation, vocabulary and basic sentences and grammar for daily dialogue (about 2 hours/week)	Copy, narrate, or describe in simple terms (around 60 words)
5 4 3	Kindergarten (3 years)	/	/

3.3.3 Data collection and analysis method

Three main types of data were collected over time: writer logs, semi-structured interviews, and text analysis. Some other documents are also collected to provide a detailed understanding of students' writing problems and difficulties, including students' discussion notes about essay writing, some written texts for various written assignments and daily communication notes with other teachers.

For written scripts, 27 essays were collected from the 15 participants. Similar to the text analysis procedure described in Chapter 2, these written scripts were also assessed in terms of the 10 discourse features using a five-point Likert scale. The first seven features were coded manually by two raters independently; the last three features were coded and re-examined

objectively by using a list of words/expressions. Inter-rater correlations of the first seven discourse categories ranged between .70 and .90, indicating good reliability. Calculating the percentage agreement (ratings that were the same or within one point difference) showed high inter-rater reliability: 95% for [F1], 95% for [F2], 97% for [F3], 99% for [F4], 99% for [F5], 94% for [F6], and 99% for [F7]. The two raters held further discussions on disagreed items and reached a consensus.

Writer log data were collected from individuals on a weekly basis. The participants shared with the researcher their writing log in a Google Doc, and some participants sent the e-copies of their logs to the researcher through email or WeChat communication. Some participants also copied to the researcher some related documents through a USB device. Instructions and sample logs were given so that they could document their processing activities with sufficient detail for the purpose of the research. The participants were allowed to write the log in either English or Chinese but were required to maintain the log from the time they received the writing task to the submission of the final paper. They were reminded three to five times every week through WeChat to record their progress in the logs. A total of 149 journal entries and around 60 hours of interview recording were collected. The writer log data were analysed in the same way as the interview data.

Based on an analysis of the participants' written scripts and writer logs, semi-structured interviews were conducted to examine individual participants' knowledge status about writing and their use of specific writing strategies. Following Dörnyei (2007), four major steps were followed in the interview data analysis: transcribing data, pre-coding and coding, growing ideas and generating conclusions. The interview recordings were first transcribed using the automatic transcribing software Xunfei, and then verified and refined through

human listening and editing. Before coding, a coding manual and a coded sample were generated, which included six typical processing activities. An experienced writing teacher was invited for double coding. The present researcher and the invited teacher coded the first two interview data (~20,000 words) independently, discussed the differences, made refinements and clarifications for each code and finalised the coding framework. The inter-coder reliability of the first two interview transcriptions reached 82%. Then, the researcher proceeded with coding the remaining transcriptions, and the second coder conducted a second coding based on the researcher's coding. A discussion was held periodically regarding any differences or disagreements. More codes/themes and subcodes were generated in the continuing process of coding (see Appendix C for details). The interview data were coded in a Word document first and then transformed into txt. format for further analysis in AntConc 3.5.8 (Antony, 2019). The functions of Concordance and File View of AntConc were used to synthesise and visualise the coding data.

3.4 Findings

This section first presents an overall profile of the participants' strengths and weaknesses, as observed from their written texts. It then reports the findings of individual processing strategies and academic writing knowledge status, as reflected in their authentic writing exercise. For each of these three tiers, related findings are presented from four perspectives, starting with the overall features across cases both in Hong Kong and in mainland, and then the six cases in Hong Kong group, and afterwards the nine cases in mainland China group, and finally followed by comparison and contrast between these two groups.

3.4.1 On Discourse features

RQ1: *What discourse problems do these EFL undergraduate student-writers display in their English essays and what are the similarities/differences regarding their discourse features if any?*

The students' discourse performance was rated against 10 DC features at five levels. A score of 5 indicated no problems or weaknesses. A score of 3–4 indicated some signs of basic development, but that special efforts were required. A score of 1–2 indicated severe problems/weaknesses.

Overall, students from these two contexts displayed basic- or fair-level discourse features (mean score = 2.97), with a mean score of 2.92 on the overall discourse feature for all participants in the Hong Kong context and 3.02 in the mainland context. The local Hong Kong participants demonstrated the best discourse features saliently among all 15 participants (mean score = 3.77). Their relative strengths were in the use of connectives ([F8] and [F9]), whereas all remaining features were weak, with a mean score approximately 3.00. Their salient weaknesses were on [F3], [F6] and [F10]. More details of the discourse features are shown in Table 3.3 below.

Table 3.3

A comparative analysis of the mean scores between mainland and Hong Kong

Category	[F1]	[F2]	[F3]	[F4]	[F5]	[F6]	[F7]	[F8]	[F9]	[F10]	Mean
Hong Kong	3.17	2.39	**1.89**	3.17	3.06	**2.61**	**3.56**	3.44	3.28	**2.61**	**2.92**
Mainland	3.67	3.22	**1.44**	3.22	3.00	**2.78**	**2.67**	3.67	3.89	**2.67**	**3.02**
Local Hong Kong	3.67	3.00	3.00	**4.00**	**4.67**	**4.33**	**5.00**	3.67	4.00	2.33	3.77
Mean	3.42	**2.80**	**1.66**	**3.19**	**3.03**	**2.70**	**3.11**	3.56	3.58	**2.64**	2.97

A shared weakness in these two contexts was related to [F10] *reader–writer interaction*, with a mean score of 2.61 in the Hong Kong context and that of 2.67 in the mainland context. There were more boosters than hedges in their writing. There was also very limited range in their use of hedges and boosters; the most frequently used type was modal verbs ('*may*', '*should*', '*will*' and '*must*'), and there were few uses of related verbs, adjectives and adverbs.

One difference between these two contexts was related to local coherence on [F7]. This feature was found to be a strength in students' writing in the Hong Kong context (mean value = 3.56) but a relative weakness in the mainland context (mean value = 2.67). All participants in the Hong Kong context could produce relatively coherent sentences within paragraph writing.

Another noticeable feature was related to the difference between students from the mainland and local Hong Kong over the features conveyed through [F3]–[F6]. These four features were a shared problem among mainland students but were relative strengths in the local Hong Kong participant writing. Take for example the participant Han in mainland case study. The whole essay read like short answers in response to a series of questions; there were no explicit linguistic features to link different paragraphs. Moreover, the beginning of the essay read like several clusters of sentences loosely hanging together, which pose a considerable challenge for a reader to understand the intention of the writing and the key context.

3.4.2 On composing strategies and academic writing knowledge

RQ2: *How do these problematic features arise in composing processes? What are the similarities/differences regarding their processing strategies if any?*

A shared feature in these two contexts was the participants' relatively sufficient research and planning activities. This reflected their commitment to disciplinary content in academic writing. There was considerable emphasis on content in disciplinary writing on the students' part; thus, the language and structure issues appeared quite diluent in the background. As stated by one mainland participant Yang, '*the disciplinary essay writing was mostly a subject task, instead of a piece of writing work in the sense of what was delivered in a writing course*'.

The second similarity was related to revision strategies. It seemed that participants in both contexts were relatively less bothered about the revision of the whole text once finishing the initial complete draft. For instance, in the Hong Kong context, Rachel finished her essay several days earlier and submitted it immediately to save more time for other courses; Tiffany finished the essay several hours earlier and then went for a weekend party; Lucia and Wei finished their initial draft minutes before the deadline and almost did not have time for revision. In the mainland context, the LP participants Yang, Han and Lin did not have the thought of revision, whereas some MPs/HPs made some efforts for revision after finishing a draft. There were at least four influencing factors. The first factor was related to L2 proficiency. The multi-case study on the mainland showed that students with low L2 proficiency often felt difficulty perceiving their own problems in writing, thus having little thought of making self-revision. The second factor was related to the skills in the monitoring of the whole composing process. Lack of awareness of the time dimension of the composing process might have resulted in a highly unbalanced time distribution over typical composing activities. As shown in the multi-case study in Hong Kong, participants Lucia and Wei spent considerable time in research and planning and, thus, had very little time left for formulation and revision. Third, one salient factor was likely to be motivation in writing. As found in the multi-case study in mainland China, the LP participants' single motivation in writing was to pass the exam, whereas completion of the initial draft often meant succeeding in passing the baseline. Thus, to these students, revision might have been unnecessary. The fourth factor was likely the means of computer writing. Computer writing had partly corrected students' grammar, spelling and other mechanics. With the conventional thought of revision on grammar, the students might have felt that there was no need for further revision.

The third similarity was lack of adequate genre knowledge, as frequently perceived from their over-reliance on model essays for planning/formulation and their confusion on differences between abstract and conclusion. Especially there are frequent mentions of model essays by mainland participants. Some participants (e.g. Yang) were so dependent on the model essay that she modelled the sample essay sentence by sentence in a very mechanical manner and with extreme exhaustion of the brain. Comparatively, the local Hong Kong participants showed good discourse features regarding global coherence. It may suggest that Chloe had some genre knowledge. However, her knowledge of the genre was not explicit, or at least not systemic, as evidenced by her comments: '*The writing course mainly teaches how to use different vocabulary, whereas my main challenge in academic writing is how to develop the topic at depth and express her opinion systematically. This is what the writing course cannot teach*'. The limited academic writing knowledge was also partially exposed in students' knowledge about related academic sources. Take Qiang in mainland for instance. When asked how he searched for relevant materials, he said he downloaded an app named 'Zhi Wang' ('知网' in Chinese). In the beginning of the interview, the interviewer took it as the short form of China National Knowledge Infrastructure(CNKI). As the interview continued, however, Qiang's understanding of 'Zhi Wang' turned out to be a short form of 'Zhi Hu Wang Zhan' ('知乎网站' in Chinese). This exposed his limited knowledge of typical academic resources and his potentially insufficient academic writing knowledge.

One salient difference between mainland participants and their Hong Kong peer Chloe was related to skills in effective coordination and monitoring of the whole composing process, especially the monitoring skills related to planning, formulation and revision. Chloe demonstrated skilled composing strategies and could create a well-written product with high-quality discourse features in the last three days before the submission deadline. Similar to the

deadline fighters who tended to prepare for the writing in the last three days before the due date, the mainland participants had considerably suffered in completing the writing assignment on time, and the evaluation of their written essays showed relatively worse discourse features compared to those exhibited by their Hong Kong peer. Despite adopting sufficient research and planning at a very early stage, the HP mainland participant Lydia achieved a lower score on discourse features than her local peer Chloe. Hence, besides the potential influence of L2 proficiency and content knowledge, the following two factors were believed to be significant contributors to the difference: rich writing experience and constant accumulation of genre knowledge. Compared to mainland students, the local Hong Kong participants had abundant writing exercises over extended writing in secondary school learning. This rich writing experience was beneficial in the consistent exercise of related composing skills and the accumulation of corresponding genre knowledge.

The second difference was related to skills in using related sources. Students in the Hong Kong context exhibited richer knowledge and use of related academic sources, such as Google Scholar, and related library resources as provided by their host university. In contrast, participants in the mainland context demonstrated poor knowledge and use of proper academic sources. They almost did not mention or use the library sources as provided by the university; instead, they tended to refer to some informal sources, such as Baidu or Bing (必应网站) for related information. One highly proficient participant (Qiang) in the mainland context even did not know CNKI, the most commonly used academic resource by mainland scholars. One influencing factor was the availability of sources. In mainland, students cannot often access Google Scholar or most websites created outside China. Long immersed in this context, the mainland students gradually lost their desire to try academic sources, such as Google Scholar or Research Gate. Another factor was related to students' perceptions of the

nature and functions of academic matters in their daily learning. In the mainland context, many students still perceived academic learning as a passive activity pushed heavily by teachers and, thus, seldom took the initiative to make good use of available sources. In contrast, in the Hong Kong context, the students were often proactive in seeking opportunities to make better use of related sources.

The third difference was related to the formulation strategy. Mainland participants, especially MP/LP students, frequently turned to machine translation for sentence or paragraph writing, whereas students in the Hong Kong context mostly wrote directly in English, even if their English was not good (Tiffany). This was partly understandable in that the LP students had difficulty writing in a foreign language; the use of machine translation could ease their worry and help them produce a written work of higher quality. An alternative explanation might be their motivation in writing. The Hong Kong participants were likely to write to both learn the content and the language, whereas the mainland participants tended to write to complete the task and gain scores.

3.4.3 A complex dynamic perspective on the similarities and differences

RQ3: *To what extent do students' processing strategies and academic writing knowledge status affect their choices and decisions during writing? What other factors may contribute to such similarities/differences?*

From the mainland case study, a positive correlation was observed between English proficiency and discourse features, with HP participants having an increasingly higher mean score of discourse features than LP participants. This is especially salient in relation to local coherence. Students' difficulty in finding appropriate vocabulary, sentence patterns and correct grammar is likely to incur two problems: coherence breaks (i.e. hampered idea progression between sentences), or production of simple/superficial ideas for the sake of safe

grammar. Table 3.4 below showed details of students' discourse features as revealed from their written essays.

Table 3.4

Students' overall discourse features

NO.	Low Proficiency				Medium Proficiency				High Proficiency				Total
	Yang	Han	Lin	Mean	Xuan	Jones	Jian	Mean	Jie	Fan	Qiang	Mean	Mean
[F1]	4	4	3	3.67	4	4	3	3.67	4	4	3	3.67	3.67
[F2]	4	3	3	3.33	4	3	4	3.67	2	4	2	2.67	3.22
[F3]	1	1	1	1.00	4	1	1	2.00	2	1	1	1.33	1.44
[F4]	4	3	2	3.00	4	3	3	3.33	4	4	2	3.33	3.22
[F5]	3	2	3	2.67	4	2	3	3.00	4	3	3	3.33	3.00
[F6]	3	2	2	2.33	2	3	1	2.00	4	5	3	4.00	2.78
[F7]	3	2	2	2.33	2	3	2	2.33	3	4	3	3.33	2.67
[F8]	3	3	3	3.00	4	4	3	3.67	4	5	4	4.33	3.67
[F9]	4	3	4	3.67	4	4	4	4.00	4	4	4	4.00	3.89
[F10]	2	3	2	2.33	3	2	3	2.67	2	4	3	3.00	2.67
Mean	3.10	2.60	2.50	2.73	3.50	2.90	2.70	3.03	3.30	3.80	2.80	3.30	3.02

Take for instance the Low-proficiency participant Yang in mainland context. Yang's specific strengths were on [F1] *topic/focus*, [F2] *thesis statement*, [F4] *reader orientation* and [F9] *connective accuracy*; her biggest weakness was on [F3] *controlling idea* and [F10] *complexity of hedges and boosters*. Meanwhile, she had some problems with [F5] *supporting paragraphs*, [F6] *conclusion*, [F7] *local coherence* and [F8] *connective complexity*. In particular, the introduction section presented some context but only in general linguistic terms, and there was no mention of the key disciplinary concepts as required in the writing prompt. In one of her supporting paragraphs, the topic sentence was not very closely related to the main ideas of its previous and following paragraphs. In her conclusion, there was clarity of her position, but the explanation seemed irrelevant to both justification of the position and analysis in previous paragraphs.

Upon the first sight of the writing task, she felt hit by a big bomb: '*Every word is a difficulty, from typical purposes of the essay, to understanding of the term problem and the concept of ethical/cultural dilemma, and to search and selection of reference*'. Despite her salient

difficulties and problems with English language, Yang was invited by her classmates to join a private eight-member group discussion on this task because of her well-acknowledged talent in idea generation.

Regarding her biggest problem with [F3], she did not present the controlling idea because her instructor did not make such a requirement. Yang's salient processing strategy was over-reliance on the sample/instructor for planning idea generation and organisation. She described most of her writing experience as an 'extended blank-filling' exercise. If a sample essay was provided, the sample was like an outline in her mind. She would imitate paragraph-by-paragraph and sentence-by-sentence by only changing some specific vocabulary related to her specific topic/focus.

Both at the formulation stage and throughout the writing process, Yang's first concern was related to finding appropriate vocabulary, sentence patterns and correct grammar.

> Yang: *It is really a torture when writing down my ideas on paper. Writing of every sentence is a torture. When writing a sentence, Gosh, how to change the tense? Later, Gosh, does the –ing form need extra changes? I am completely confused. The situation is that every time I write down a sentence, I am not confident. I think the sentence is problematic.*

In the formulation stage, Yang's English language proficiency was insufficient to express her ideas. Her choice was to change her ideas based on the availability of her vocabulary, sentence and grammar knowledge. Toss and turn with vocabulary and grammar had taken most of her energy in the formulation stage that she could hardly have time or emerge to

think about other subtle features, such as the delicacy of *reader–writer interaction*. Meanwhile, Yang made no revision. One of her reasons was that time and space limitations did not allow structural revisions. The other reason was that she could not recognise her own vocabulary/grammar errors/problems, nor would she have the ability to revise even if she could find the problems. Hence, composing the draft was mostly Yang's first-and-final product of the essay task.

The influence of L2 proficiency on discourse features were also perceived from medium proficiency participants. Take Jones in mainland for instance. Jones had problems with [F5] supporting paragraphs. In particular, it was difficult to identify a proper topic sentence in two of her supporting paragraphs. When talking about topic sentences, she said, '*Every paragraph needs one sentence to convey the main point. But when wanting to express myself in one sentence, my ideas are incomplete. And I also want to keep the topic sentences in consistent sentence patterns. Thus, I struggle a lot at selection of the key words*'. She also shared that she cared much about grades and worried that she would lose part of the grades if she made mistakes. This indicated that she had good knowledge about the features and functions of topic sentences, but her worries about grades stopped her from demonstrating her best knowledge and skills, and she only turned to the parts that she was most sure about, to secure the grades. Take another example Jian in mainland case study. Jian had problems with supporting paragraphs. In particular, one topic sentence in Jian's supporting paragraphs was not clear: '*Next, the problem solutions should be analyzed*.' When asked what he meant by '*problem solutions*', he said it refers to a set of principles and strategies for problem-solving issues and he took '*problem solutions*' as one concept. This indicates that he had a rich idea but had difficulty expressing himself in written English. He described his English as simplistic and inadequate ('简陋'). There would be frequent mistakes if he wrote complex or

compound sentences. Hence, inaccuracy of vocabulary and sentence grammar is an important factor contributing to his organisation of supporting paragraphs.

Another noticeable interaction is the local coherence and students' use of machine translation. Machine translation both helps and hampers students' writing effect. Still take Han in mainland case study for instance. One day before the formal exam, Han wrote down three complete drafts for the three topics, because she worried that she could not convey her ideas clearly in English in formal class writing. She first wrote each draft in Chinese and then translated her drafts into English through online translation websites. Afterwards, she recited one essay, which later happened to be an exact hit on her exam topic. On the exam day, she wrote her essay by reciting her draft from memory. Han read mostly in Chinese because she had difficulty reading academic materials in English. She would translate a whole sentence into English by putting it into an online translation website, and then directly adopted the automatic translation in her essay.

> '*To control emotion with reason requires the leaders to have a correct social cognition and value or intention.' (Shujun, 2013). <coherence break> What should coach Jeff do as the leader of this basketball team? There are two plans here…*
>
> ---Excerpts from one body paragraph in Han's essay

The translation strategy could incur textual problems, such as confusing paragraphing on [F5] and coherence breaks within the body paragraph on [F7].

Meanwhile, besides machine translation, local coherence in student writing was found closely related to citation skills. Take Xuan in mainland for example. Her salient weakness was in

writing the conclusion. Her conclusion included the writer's position and some justifications. However, the explanations were unclear and mainly included citations and integration of references, with no specific references to previous paragraphs. In the interview, Xuan expressed her mechanical addition of references for fulfilling the reference number requirements in writing the conclusion.

Xuan: *For the last paragraph, I was talking about nonsense.*

Researcher: *It seemed you did not finish writing the essay…*

Xuan: *It's finished. I wanted to add more, but I was worried it might go beyond the required length. In the conclusion, I forced myself to integrate two cases of citation...*

Researcher: *Reference does not mean direct quotations, but you wrote about four lines of direct quotation….*

Xuan: *Yes. Otherwise, I had to paraphrase the language. But if I paraphrased them in my own words, I worried about my mistakes. Because the sources I took notes of did not fit the body paragraphs, I have to use them in more general points like the conclusion part.*

Hence, in writing the conclusion, Xuan considered covering mostly what was mandatory, as required by her instructor and the writing prompt, and the requirement on reference numbers facilitated her mechanical addition of largely irrelevant sources of information.

One remarkable difference was related to the motivation of academic reading. Participants in the Hong Kong context had a strong motivation to learn the related knowledge. Thus, they carefully read the related articles and other learning materials. They spent considerable time reading the literature directly in English and making detailed notes, even if without subject

teachers' requirements. In contrast, in the mainland context, most participants had the habit of camera reading (i.e. using the whole-page camera translation function to translate English into their mother tongue) for reading academic materials written in English. One primary intention of these mainland participants' reading was to find helpful information for fulfilling the citation requirement rather than having an accurate understanding or synthesis of related ideas. Take Jian for instance. When asked what a good essay is, Jian responded, '*An essay with the highest grades is a good one. How to achieve high grades? Meet the teacher's requirements! If you could tick every box on the teacher's list of writing requirement, it is a good essay*'. To pass the exam and achieve good grades, the students' core concern did not focus on what was indeed an effective essay and why they should write this way. Rather, their concern was to try to fulfil whatever their instructor's requirements were and to avoid any potential risk or uncertain move that might incur a loss of grades. Such motivation leads to frequent mechanical addition of references just to fulfil the reference number requirement, which subsequently incurred coherence problems.

Relating to mainland participant, an accompanying phenomenon with students' score driven motivation was their reluctancy to think or being inert in thinking on their own. When clueless about the organisation, their dominant processing strategy was to rely on a model essay/instructor for planning idea generation and organisation. As there was no explicit requirement on the presence of a controlling idea, there was an understandably absence of a controlling idea. They reluctancy to think may be also revealed from their *question-answer planning/formulation strategy*, *whole-page automatic translation in reading English academic materials, and* mechanical addition of references for fulfilling the referencing number requirement.

Quite on the contrary, students in Hong Kong context displayed strong motivation and sincere thinking in completing their writing assignments. Take Chloe in Hong Kong for instance. She considered "interest" and "relatedness to disciplinary study" as the two critical elements in her decision or selection on a writing topic and focus. When the provided topics did not match her interests, she would negotiate an alternative topic. In contrast, the mainland students would choose from the given topics or compromise with teachers' requirements when the given topics were not satisfying; they did not show very strong interest in certain topics and did not negotiate with instructors on specific topic decisions.

One potential influential factor on learner attitude and motivation seems to be related to L2 proficiency. Besides the majority participants' inert thinking and score/-led decision, there are also participants seeking for personal best and enjoyment in writing. Take for instance the high-proficiency participant Jie in mainland China. She took a proactive attitude toward learning, not worrying about whether she would pass the exam; she just wanted to deliver her best performance.

3.5 Discussion

Academic writing is a complex dynamic process that occurs overtime and across multiple situations. The present study explored student writers' academic writing at three levels. As stated in Lasen-Freeman (2017), "what is important in a complex system is the interdependent relationship of the factors that comprise it" (p.27). Guided by the key principles of complexity theory, this section further discussed the inter-connection and interaction within and among specific discourse features, processing strategies and academic writing knowledge.

Immediate attention concerned [F3] *controlling idea*. Most of the essays that we analysed had no controlling sentences. There are two potential underlying factors for this: insufficient academic writing genre knowledge and over-reliance on modal essay/instructor for planning. First, the writing instruction/guide had no explicit and mandatory requirement regarding the presence of a controlling sentence that indicated the overall essay map. The students seemed to be passive followers instead of active agents in communicating the overall intention. The central goal of developing writing skills is to gain executive control over the composing processes so that one can respond adaptively to the specific needs of the task at hand and enable knowledge creation or transformation (Kellogg, 2008). To achieve this goal, students' concerns must be shifted from the final written product to the specific composing processes. Particularly, among the six typical processing strategies, enhancing students' self-evaluation strategies and skills in monitoring and controlling the whole composing process might be especially helpful in building student agency in writing. As this evaluation mainly engages individuals' appraisal of the utility and effectiveness of the learning product and process (Wenden, 1998), for developing an evaluation strategy in the present research context, the first step is to ensure the students' proper understanding of the criteria or standard of effective writing. Based on the clarification of the proper criteria, skills in self-monitoring and control enable students to observe their own current thoughts/behaviours in writing and make modifications according to specific contexts instead of rigidly following the model essays or teacher instructions.

Global coherence is a frequent problem for all three levels of students. Global coherence is closely related to genre knowledge, reasoning and argumentation and disciplinary knowledge. In students' daily disciplinary writing exercise, the genre features and key reasoning patterns are often prescribed in the subject teacher's writing guide and further informed through

students' constant consultation with their instructors. However, such knowledge remains largely invisible in students' narration of their disciplinary writing experience. This indicates that they do not have sufficient acquisition of genre knowledge and/or cannot transfer what they learn from English writing classes to subject learning. Furthermore, the insufficiency of genre knowledge is also reflected in their planning strategies in idea organisation, especially in the case of MP and LP participants. These students are over-reliant on exemplars and teacher instruction for conceiving the detailed outline of the essay. Thus, they tend to make, in Swales' (1990) term, a parody or a worse travesty and burlesque of their writing. Hence, instructions on a flexible application of genre knowledge and reasoning and argumentation patterns are necessary to help these EFL learners transfer global coherence across courses. Based on the distinction between genre knowledge and genre awareness, Tardy et al. (2020) proposed a comprehensive theoretical framework on genre knowledge and provided further guidance of genre knowledge instruction in a classroom setting. Given students' potential insufficiency in genre knowledge and need to transfer skills across courses, the present study suggests that the first step is to heighten students' genre awareness.

Local coherence [F7] is another common problem. Coherence breaks often occur when students add references, and unrelated idea progression frequently occurs when students try to relate theory to their own analyses. Irrespective of whether there is a real connection conceptually, the linguistic expression does not convey the potential connections between these diagnosed problematic sentences. Coherence breaks are directly caused by two processing strategies: mechanical addition of references and linguistic difficulties in finding appropriate vocabulary, sentence pattern and correct grammar expressing themselves at the formulation/revision stage. These strategies are also closely related to their strategy to read source materials in grossly translated forms and a surface approach to read for fulfilling the

referencing requirement at the research stage. Research is essential in academic writing and provides primary sources for problem formulation and idea generation. Meanwhile, for novice English L2 writers, reading English academic materials scaffold an effective platform for improving their academic language proficiency. However, as observed by Hamilton (2018), the significance of academic reading in appropriate academic writing is not fully understood. Hence, an early and guided learning of academic reading is necessary. Given the intended audience and difficulty level of target academic reading materials, a starting point for novice L2 students might be textbook materials, lecture slides, other book chapters, related academic websites, dissertations and the more specialised and complex peer-reviewed articles at the endpoint.

Furthermore, there are two processing strategies adopted by the participants that may be worth further exploration: machine translation for reading/writing and blank-filling and question/answer strategy adopted for idea generation/organisation. The impact of machine translation on writing is controversial in L2 writing studies, having both drawbacks and benefits (Lee, 2020). On one hand, it can facilitate better communication of ideas with less effort (Garcia & Pena, 2011) and enhance learner confidence and motivation (Lee, 2020); on the other hand, it can incur laziness in learning (Garcia & Pena, 2011), as well as problems of accuracy, fluency and appropriateness (Lee, 2020). Lee (2021) stated that machine translation can work effectively with proper teacher/peer feedback. In contrast, in the present study, the participants mostly worked with machine translation individually, with no efforts in seeking peer/teacher feedback. Similar findings have been obtained regarding the use of machine translation for academic reading as well (Garcia & Pena, 2011). Hence, there is a need for proper teacher guidance on the use of machine translation for effective writing/reading, particularly for beginners or intermediate-proficiency learners.

Another noticeable composing behaviour was the blank-filling or question/answer strategy in idea generation/organisation. This strategy reflected students' limited knowledge on related genres. This is in line with the observations of EFL learners' two-dimensional perception of academic writing in many studies (i.e., content, knowledge and linguistic skills) (Leki, 2007; Ruan, 2014; Zhao & Lyv, 2019). The role of model text or the use of templates has been frequently discussed in L2 writing studies (Hyland, 2004; Johns, 2011; Peloghitis & Ferreira, 2018). There are considerable benefits in using model essays as a pedagogical tool: raising awareness of organisational conventions (Hyland, 2004), creating a mental model of rhetorical moves (Crinon & Legros, 2002) and easing writer anxiety (Macbeth, 2010). While embracing the multiple benefits of model essays (exemplars), this study found that over-reliance on model texts might give students the impression that a rigid format is required for a specific assignment, instead of the acquisition of flexible genre knowledge potentially transferrable across task types. Adequate instruction on genre learning can help address the structure concerns in L2 writing and make novice L2 writers feel that there is space to critique and change when referring to the models.

3.6 Conclusion

Overall, the participants demonstrated a basic level of discourse features. Their relative strengths were on *topic/focus*, *connective complexity* and *connective accuracy*; their weaknesses were on *controlling idea*, *conclusion*, local coherence and *reader–writer interaction*. At the middle were the thesis statement, reader orientation and supporting paragraphs. There was a positive correlation between English proficiency and discourse features, with the HP participants having increasingly higher mean scores on the discourse features than the LP participants. In addition, the HP students were distinguished from their

MP/LP peers with good academic reading habits and proactive attitudes towards learning. A common reading strategy adopted by the LP/MP students was *whole-page camera translation in reading English academic materials*. For the LP students, another typical processing strategy was *question-answer or blank filling for planning/formulation*. Furthermore, the comparative analysis in the two contexts of Hong Kong and mainland China identified more differences than similarities. Three salient similarities were their sufficient research/planning, insufficient revision and inadequate genre knowledge. Four major differences were monitoring of the whole composing process skills in using related sources, formulation strategies and motivation in academic reading. Attitude and motivation were also found to be critical contributing factors. In response to the identified problems and difficulties, incorporation of a series of purposeful academic reading activities might work as a meaningful treatment measure in EFL learners' writing courses.

Nevertheless, the participants sampled in this study, especially in the case of Hong Kong study, were restricted due to limitation of the recruitment criteria and the disciplinary writing contexts explored in this study may be also not representative. Thus, interpretation and implications of the findings need to be taken with care for generalization of the results to the wider population. In next chapter, a quantitative survey study was initiated for the purpose of identifying potential trend of the target EFL learners writing problems and difficulties at a more generalizable degree.

Chapter 4 Study Three — A Survey Study

Chapter Abstract

The survey study investigated the knowledge and composition strategies of a group of 277 undergraduates writing academic essays in an international joint degree program at a university in mainland China. The study found that a five-factor processing model and a four-factor knowledge model can explain students' processing strategies and academic writing knowledge. Further analysis showed that the students on average exhibited basic processing strategies and academic writing knowledge. There were significant differences regarding overall processing and revision between high- and low-level student writings, but no significant differences were found across three L2 proficiency levels. The content analysis of the open-ended questions yielded consistent but more detailed findings. The top three most frequently reported student strengths were idea organisation, logic in argumentation and research of related sources. The top three reported weaknesses were lexical complexity, grammar and disciplinary background knowledge. Particularly, more than half of the students (58%) expressed their need to increase their vocabulary related to their disciplinary studies.

4.1 Introduction

English academic writing is crucial to academic success in university studies, especially in English as a medium of instruction (EMI) contexts (Nesi & Gardner, 2012). Apart from universities in English-speaking countries, EMI is practised in many post-colonial countries and regions, such as Hong Kong, and even in non-English-speaking countries (Lasagabaster et al., 2021). In mainland China, there is an unprecedented increase in university EMI programs in the past decade. These programs often involve cooperation with overseas universities and employ English as the primary medium of instruction. Considering the differences between secondary school learning and university study and those between school

learning in China and in many English-speaking countries, academic writing likely poses particular challenges to novice EFL learners in these contexts (Evans & Green, 2007).

This chapter presents findings from a survey study of 277 undergraduate students involved in a Sino-Australian cooperative education program. The survey focused on students' academic writing knowledge, the processing strategies they adopted and the discourse competence features in their academic essays. Individual writing problems and difficulties were also investigated through complementary open-ended written questions at the end of the questionnaire.

4.2 Literature review

This section first reviews typical teaching and learning approaches adopted in classroom writing during university studies. It then reviews empirical studies on EFL learners' composition strategies and their perceptions of their own academic writing knowledge and writing practices. At the end, specific research focuses and methods adopted by the present study are presented.

4.2.1 Teaching and learning of academic writing in EFL learning contexts

There are currently two main approaches to academic writing instruction in higher education: the English for specific purposes (ESP) approach and the academic literacies (AL) approach (Coffin & Donohue, 2012; Flowerdew, 2020; Gardner, 2012; Hyland, 2019; Swales, 2012; Wingate, 2018; Xu, 2015). These two approaches have different theoretical underpinnings and research focuses. The ESP approach is rooted in language teaching, has its theoretical underpinnings in the systemic functional linguistics theory and focuses on non-native speakers (NNS) of English. The primary concern of the ESP approach is with text, focusing

on language in use and the language system in academic writing (Ding & Bruce, 2017; Pérez-Llantada & Swales, 2016). On the other hand, the AL perspective has its theoretical underpinnings in the new literacy school; it is less concerned with the qualities of successful texts and more with practitioners' and learners' writing experiences in context (Barton, 1994; Lillis & Scott, 2007; B.V. Street & B.B. Street, 1984).

Academic writing instruction adopting the ESP model has developed clearly defined theoretical underpinnings and established teaching methodologies (Ding & Bruce, 2017; Pérez-Llantada & Swales, 2016; Wingate, 2012). The most influential method is the genre approach to academic writing. The seminal work is Swales' (1990) proposal of the genre-based approach to the understanding, teaching and learning of English in academic and research settings. In his work, a genre is understood as 'a class of communicative events, the members of which share some set of communicative purposes' (Swales, 1990, p. 58). Purpose, rather than similarity of form or other criteria, is considered the primary determinant of genre membership. Askehave and Swales (2001) further suggested a five-step text-driven procedure for genre analysis, starting with the analysis of structure, style, content and purpose and followed by the analysis of genre, context, repurposing of genre and finally reviewing genre status. Typical of this approach is genre analysis, in which students are asked to analyse the genres in which they will write. In addition, corpus tools are sometimes incorporated to allow the analysis of a more extensive set of texts and to generate greater generalisability of findings.

The latest development in academic writing pedagogy is the process-genre approach in the 2000s. Badger and White (2000) were among the first to propose the process-genre instruction that relates textual features, cognitive processes and contexts. Badger and White

(2000) conceived writing as ‘a series of stages leading from a specific situation to a text, with teachers facilitating learners’ progress by enabling appropriate input of knowledge and skills’ (p. 156). Moreover, they believe that different genres require different sets of knowledge and skills, and the writing development varies for different groups of learners. Their model engaged two major moves: (1) starting with learner orientation within a particular social situation and proceeding to (2) scaffolding with exercises of processing strategies, such as planning, drafting and proofreading. This approach is claimed to be effective in facilitating both L1 and L2 writing. But there are still few empirical reports (Huang & Zhang, 2020) clarifying concrete procedures and the effectiveness of their application to practical teaching and learning.

Specifically related to the teaching of writing to L2 EFL learners, there are two further main dimensions/perspectives: learn to write (the LW dimension) and write to learn (the WL dimension). In terms of the WL dimension, there are sometimes further distinctions between writing to learn content (WLC) and writing to learn language (WLL). Manchon (2011) noted that the LW dimension originates from L1 composition, EAP and the Writing Across the Curriculum movement in America. The WL perspective, however, is closely related to the pedagogical perspectives of content-based instruction (CBI) and content and language integrated learning (CLIL) in the European contexts (Hirvela, 2011; Mohan et al., 2010). A typical practice in the WL dimension is writing in the disciplines, through which students ‘display their critical and analytical skills, their use of English for reasoning and persuasion, their grasp of subject matter issues, and their ability to shape an argument using the conventions of their field’ (Hyland, 2013, p. 241). In the context of L2 academic writing in a local non-English setting, both L2 proficiency and disciplinary content are found to pose significant difficulties and problems in daily learning (Ruan & Chen, 2017; Zhao, 2019b).

Especially in the EMI context in L2 writing, students must learn subject content and the English language simultaneously (Galloway & Rose, 2021). Thus, the CLIL pedagogy from the WL perspective may better serve novice L2 writers' needs in university studies.

4.2.2 EFL learners' processing strategies and existing research methods

Process in writing refers to the activities the writer engages in to produce texts (Prior, 2004). From the cognitive perspective, process equals strategies that are primarily focused on producing a final product that achieves a writing goal. As pointed out in Susser (1994) and Stapleton (2010), academic writing is a recursive, responsive process that takes place over time on multiple occasions and in multiple settings. Given the dynamic and often complex social-cultural and situational contexts in which writing takes place, the investigation of writers' processing strategies has been conducted through diverse methods. The main methods include think-aloud protocols (Roca de Larios et al., 2008), retrospective interviews (Green, 2013; Stapleton, 2010), video or audio recordings (Green, 2013), keystroke logging (Baaijen & Galbraith, 2018) and questionnaires (Qin & Zhang, 2019; Zhao & Liao, 2021).

Think-aloud is a common method used to understand learners' composition process (Beauvais et al., 2011; Roca de Larios et al., 2008). Using the think-aloud method and focusing on L1 Spanish speakers learning English as their L2 English in timed argumentative writing, Roca de Larios et al. (2008) found that formulation took up the largest portion of composition time, and writers' L2 proficiency significantly influenced the time allotment across different activities, with high-proficiency writers showing more balanced time distribution than low-proficiency ones. Using the same method, Beauvais et al. (2011) investigated the interrelation between students' processing strategies and the quality of narrative and argumentative texts. They found that students adopted different processing

strategies for these two types of writing, and better texts were associated with goals that emphasized explicitly the quality of the assessment criteria. Planning and appropriate, deliberate organising of ideas are considered particularly crucial in argumentative writing. However, the think-aloud method has often been considered an intrusion on the composition process.

Retrospective interviews are another popular method for understanding the composition process. For instance, Stapleton (2010) explored the composition process of an L2 learner writing a 4,000-word English essay with interviews, a questionnaire and in-depth writer logs. He identified notable differences in time allotment across different writing processes; the participant in his study allocated considerable time to research and much less time to formulation, revision and evaluation. Using semi-structured interviews and audio logs, Green (2013) conducted a longitudinal case study and examined the composition process in completing related course assignments by three EFL learners studying in a British university. He identified two approaches to writing: a planning approach (with dominant efforts in detailed planning) and a drafting approach (with dominant efforts in drafting). He concluded that perhaps the key to quality writing is not the use of a particular approach but the level of the writer's investment in a particular approach. Nevertheless, interviewing has limitations in that the writers' recollections of the writing process may not be accurate, and some details may be forgotten or described in a vague manner.

Some researchers have adopted keystroke logging to research the underlying process of writing (Bowen & Thomas, 2020; Choi & Deane, 2021). For instance, Choi and Deane (2021) used keystroke logging to analyse a group of 798 adult EFL learners who completed summary and discussion writing tasks. They identified five factors that could account for the

correlation between stable process features and overall writing quality: the amount of writing, writing fluency, variance in pause durations, fluency characteristics and pauses in writing. The quality of L2 writing was also found to correlate with keyboarding skills. Although it is beneficial to automatically analyse processing patterns and unobtrusively capture the dynamic features of typing and revision, it is challenging to interpret these activities in relation to cognitive processes (Wenglin et al., 2019). In addition, keystroke logging may be better suited to research on short, timed writing than to research on extended academic writing completed over time on multiple occasions and in multiple settings.

There have also been survey studies on students' composition strategies (Qin & Zhang, 2019; Teng & Zhang, 2016; Zhao & Liao, 2021). For example, Teng and Zhang (2016) developed a self-regulation questionnaire and validated its use in EFL writing. Their confirmatory factor analysis identified nine factors representing L2 writing strategies: text processing (TP), course memory (CM), idea planning (IP), goal-oriented monitoring and evaluating (GME), peer learning (PL), feedback handling (FH), interest enhancement (IH), motivational self-talk (MT) and emotional control (EC), with self-regulation as an overarching construct encompassing all of these other strategies. Among these strategies, six had significant predictive effects on L2 writing quality (all except IE, CM and PL). Also adopting a post-test questionnaire approach, Zhao and Liao (2021) examined the metacognitive strategies of 200 EFL students who had completed a university placement writing test. Their exploratory factor analysis identified five main types of strategy: task interpretation, planning, translating, evaluating and monitoring, and revising. However, their study observed a limited mixed correlation between the use of these strategies and the quality of the final written product. Nonetheless, despite its broader generalisability, the data elicited through a questionnaire may be superficial and may not reveal fine-tuned insights into underlying factors.

4.2.3 EFL learners' perceptions of their own academic writing knowledge

Academic writing knowledge encompasses the range of knowledge and skills that students must acquire to communicate competently in academic writing. These forms of knowledge and skills may be examined first-hand by analysing written texts and the writing process. They may also be perceived by eliciting responses directly from teachers, students or other related stakeholders. These forms of knowledge and skills, as demonstrated in the written texts, are examined in Study 1, Chapter 2. Apart from textual analysis, two other typical methods (survey and interviews) were used to directly elicit students' academic writing knowledge.

Surveys are a popular approach to evaluating students' writing difficulties (Evans & Green, 2007; Evans & Morrison, 2010; Shepard & Morrison, 2021; Zhao & Lyu, 2019). In general, among the four macro-language skills (i.e., reading, writing, listening and speaking), writing has frequently been identified as the most difficult in university studies (Evans & Green, 2007; Evans & Morrison, 2010; Shepard & Morrison, 2021). For instance, Shepard and Morrison (2021) conducted a survey of 636 first-year undergraduates in Hong Kong and found that of the 15 investigated items, the top three challenges were using appropriate academic style, planning written assignments and expressing ideas in correct English. Students also had problems with proofreading written work, writing the different parts of the paper, linking sentences smoothly, writing reference sections, integrating sources and so on. Based on Snow and Uccelli's (2009) four-dimensional model, Zhao and Lyu (2019) conducted a survey to examine 177 Chinese undergraduate English major students' metacognitive knowledge levels. They found these students conceived academic writing mainly in two general dimensions: content knowledge and language skills, with an additional two largely invisible dimensions (genre mastery and reasoning/argumentative strategies).

Interviews are another popular approach to understanding students' academic writing knowledge (Pessoa et al., 2014; Ruan, 2014; Ruan & Chen, 2017). Pessoa et al (2014) conducted a four-year longitudinal study of 23 EFL learners' writing challenges and development in an EMI university in Quatar. They found students were particularly challenged in their vocabulary knowledge for academic reading and in understanding the genre and style requirements of writing tasks. Students also struggled with time management, mainly due to limited reading and writing experience. Ruan (2014) explored Chinese EFL learners' metacognitive awareness within the broader domain of cognitive writing strategies through group interviews. He found that priority was given to language forms (lexical variety and rich diction) and content. Ruan and Chen (2017) conducted a situated qualitative study to understand students' experiences of L2 disciplinary writing at an English-medium university in mainland China. Their study generated four main themes in understanding L2 disciplinary writing: (1) language use in L2 disciplinary writing, (2) perceptions of rhetorical norms and disciplinary conventions, (3) valued practices in learning to write in the disciplines and (4) the role of writing in developing disciplinary competencies. They found that students in their final year still considered grammar and vocabulary the main challenges in their writing. Moreover, they often depended on translation to integrate knowledge from Chinese sources and construct knowledge in written English. Students also lacked sufficient discipline-specific knowledge about genre conventions.

4.3 Research design and methods

The survey study was designed to address RQ4: *What academic writing knowledge do students have, and what typical processing strategies do they adopt in academic writing?* Three sub-questions were further designed:

RQ 4a. *What typical processing strategies and academic writing knowledge do students exhibit in academic essay writing?*

RQ 4b. *What are their relative strengths and typical difficulties in the composition process and with their academic writing knowledge?*

RQ 4c. *What are the similarities/differences among individual processing strategies at three levels of English proficiency?*

4.3.1 The teaching and learning context

The survey study investigated the academic writing knowledge and the processing strategies of a group of 277 undergraduates writing academic essays in an international joint degree program at a university in mainland China. In this program, the students were expected to compose academic essays in English for both language courses and disciplinary courses. This survey was based on students' 500-word essays written in a disciplinary course (i.e., *International Business Communication*). The 500-word essay writing task asked students to apply a series of problem-solving principles and techniques to analysis of a business scenario. There was also a requirement on the number of references, i.e., 4~5 sources. The assessment requirement of this task was introduced to students at the very beginning of semester learning. Related reading materials were released to students one week before their formal writing of this essay.

4.3.2 The focal participants

The survey participants were year-two undergraduate students studying in an international education program in a mainland university. After the data collection, initial screening and further data cleaning, a total of 277 responses were retained for the final study. They were all

majoring in business studies. The student ages ranged from 18 to 21, and there were 90 males and 187 females.

4.3.3 Questionnaire

The questionnaire was first developed based on Stapleton's (2010) theoretical framework of processing strategies and writing strategy surveys developed by Qin and Zhang (2019) and Zhao (2019b) to evaluate academic writing knowledge. A second version of the questionnaire was developed by integrating task-specific writing instructions and exploratory qualitative study findings, as discussed in Chapter 3. Some features were also integrated based on the theoretical conceptualisation of the composition process and findings from related empirical studies, as reviewed in Chapter 3. Expert advice was then sought regarding the theoretical rationale, construct validity, item clarity and readability. Details of the survey were shown in Appendix E.

The survey was piloted with four groups of students. The first group consisted of four participants; they tried the survey based on one-to-one meetings with the researcher. Three other groups piloted the survey afterwards. Four main aspects were given special attention: the duration (whether there are too many items), the clarity of each item (whether there is ambiguous wording), difficulty level (whether each item is too difficult to respond to) and redundant items (which may be eliminated because they do not provide any unique information or because they measure something irrelevant). In one pilot study, the participant suggested a distinction between task instruction (题目要求) and writing guidance (写作要求). Task instruction refers to the instructions provided to complete a specific task, whereas writing guidance refers to a teacher's guidance on general writing procedures and assessment. The term 'essay 的题目要求' was changed to '题目中的写作要求' (*task instructions as*

prescribed in this essay). An in-person classroom survey was adopted rather than an online survey because it was suggested that students are frequently invited to complete surveys online, and they may give unthoughtful answers to online surveys.

Table 4.1

Specific items for collaboration strategies and linguistic skills

NO.	Item	Target construct
1	*If I do not understand the task requirements or I experience task-related issues, I will consult teachers or discuss with my peers.*	Collaboration
2	*I will seek teacher consultation or peer help if I encounter difficulties in preparation for writing.*	Collaboration
3	*If possible, I will seek feedback from teachers or peers after the first draft.*	Collaboration
4	*I can use English correctly, including grammar, spelling and punctuation.*	Language skill
5	*I have an adequate range of disciplinary vocabulary.*	Language skill
6	*I can distinguish between spoken and written English.*	Language skill
7	*I can write fluently in English.*	Language skill
8	*I can make self-evaluations and revisions concerning grammar correctness, word choice, spelling and punctuation.*	Language skill

A final version was completed based on suggestions from the subject and language teachers and the pilot study results. The final survey consists of 34 items for processing strategies, 19 items for academic writing knowledge and three open-ended questions. All items are rated on a six-point Likert scale (1: not at all to 6: very much). The items for collaboration strategies and linguistic skills are shown in Table 4.1 above for a more detailed illustration of the key constructs. In Table 4.1, collaboration is represented by three items (NO. 1, 2, and 3) and language skills by five items (NO. 4, 5, 6, 7 and 8). A higher score on Item 1 (e.g., 5) indicates that the participant was more skilled at this type of strategy; a lower score on Item 4 (e.g., 2) suggests that the participant has weaker language skills.

4.3.4 Data collection and analysis methods

The survey was distributed to the target students in paper form in their English language classes the week after students finished writing their 500-word discipline-specific essay. A total of 329 students were invited to participate in the survey. The average duration of survey completion was about 20 minutes. A total of 298 students (91%) returned their survey responses. All the responses were examined to filter out possible invalid answers, including incomplete and undifferentiating answers. A total of 277 (84%) valid responses were attained.

This set of data was then logged in an Excel file and stored in a computer, and each survey was assigned a numerical identification code. Students were grouped into three proficiency levels based on their Gaokao English results and one disciplinary writing result respectively. On one hand, students were grouped into three English proficiency levels based on their Gaokao English results for examining possible correlation between English language proficiency and their English academic writing strategies. Gaokao English test was a national college entrance exam, the score of which was attained by multiple grading of two or three experienced evaluators, as required by the grading rules. Thus the score had high validity and reliability. Meanwhile, the present research was conducted one and a half year's later after students' GaoKao exam. Thus, it was assumed that the GaoKao result may well represent students' English proficiency level within this 1.5-year time span. On the other hand, students were grouped into three English writing proficiency levels based on their disciplinary writing results as evaluated by their subject teacher for examining possible correlation between the composing strategies and writing effectiveness. Their subject teacher holds a doctoral degree and had about 5 years' learning experience in an English-speaking country. The subject teacher was also the only one who had kept evaluating similar essay tasks over the past three years for all the students on the current research site. Given the subject teacher's experience

and expertise, the evaluation results on the disciplinary writing were assumed to have reasonably acceptable reliability and validity.

Students' English proficiency was coded into three levels: 1 (LP, or low English proficiency) if their English GaoKao (national college entrance exam) score was <120, 2 if between 120 and 129 (MP, or medium English proficiency) and 3 if ≥130 (HP, or high English proficiency). This procedure resulted in 46 participants being classified into the HP group (17% of the sample), 122 into the MP group (44%) and 109 into the LP group (39%) in terms of L2 proficiency. Students' writing proficiency was also coded into three levels: 1 (LP, or low English proficiency) if their writing was scored at 14 or below, 2 if scored at 15 and 3 if scored at 16 or more. This procedure resulted in 46 participants being classified into the LP group, 120 into the MP group and 107 into the HP group in terms of writing proficiency.

Five main data analysis methods were employed: principal component analysis (PCA), confirmatory factor analysis (CFA), descriptive statistics with one-way ANOVA, content analysis and distribution frequency analysis. For PCA and CFA, the data set was first split in half with consideration on balanced proportion of students' proficiency levels, resulting in a total of 139 cases in the first half and 138 in the second half. The first half was subjected to PCA through IBM SPSS Statistics version 26 to explore the underlying constructs; CFA was conducted with the remaining half of the data through IBM AMOS version 26 to find how well the items measured the latent factors as depicted in the conceptual model. Afterwards, the resultant variables were subjected to one-way ANOVA analysis to obtain a descriptive analysis of the overall writing situation.

Regarding content analysis and distribution analysis of the qualitative data generated from the open-ended questions, the researcher first analysed the questions using a set of codes (as shown in Appendix C), and distribution frequencies were then computed overall and at each of the three proficiency levels. The responses to open-ended questions were coded twice in two weeks to sustain higher validity of the coding results. Statistics from a random selection of 100 participants showed that the overlapping percentage of the two rounds of coding was 89% for questions (a) and (b) and 91% for question (c). The final codes were achieved from a synthesis of two rounds of coding results. The distribution percentages within each group of proficiency levels were further computed and compared across students of different proficiency levels. The following section reports the details of the findings.

4.4 Findings

This section focuses on students' perceptions of their own processing strategies and academic writing knowledge as found in the survey study. First, the typical processing strategies and academic writing knowledge levels are reported. Then, students' relative strengths and typical difficulties/problems are reported. Finally, similarities and differences across three English proficiency levels are reported.

4.4.1 The overall features

RQ 4a. *What typical processing strategies and academic writing knowledge do students exhibit in academic essay writing?*

An analysis of the survey data on processing strategies was first conducted. The assumptions of principal component analysis were first tested for their suitability for principal component analysis. An initial EFA extracted six factors from 27 items that might explain students'

typical composition strategies. The six components all had eigenvalues greater than one and explained 32.9% (research), 11.6% (planning), 6.3% (formulation), 6.2% (revision), 5.1% (evaluation) and 4.0% (collaboration) of the total variance, respectively. In addition, in the six-factor solution, the commonalities of all the items were above .5, and the six factors could explain 66.07% of the total variance, meeting both the solution criteria and the interpretability criteria.

After exploring the potential factor structure with PCA, a check was performed to verify whether the data set fit the identified structure and the theoretical model. The first round of CFA for the initial 27-item, six-factor structure did not show a very good model fit. Based on a correlation analysis, three items on one factor (collaboration) were removed, and the remaining 24 items were subjected to the second round of CFA. According to Browne and Cudeck (1992) and Kline (2015), the results of this round of CFA showed a reasonably acceptable model fit ($\chi 2/df = 3.07$, TLI = .84, CFI = .86, RMSEA = .09, SRMR = .08). The internal reliability of each of the five factors was also computed. Cronbach's alpha reached .88 for research, .82 for planning, .85 for formulation, .86 for revision and .78 for evaluation, indicating high reliability. As a result, the 24-item, five-facet structure was retained for final analysis.

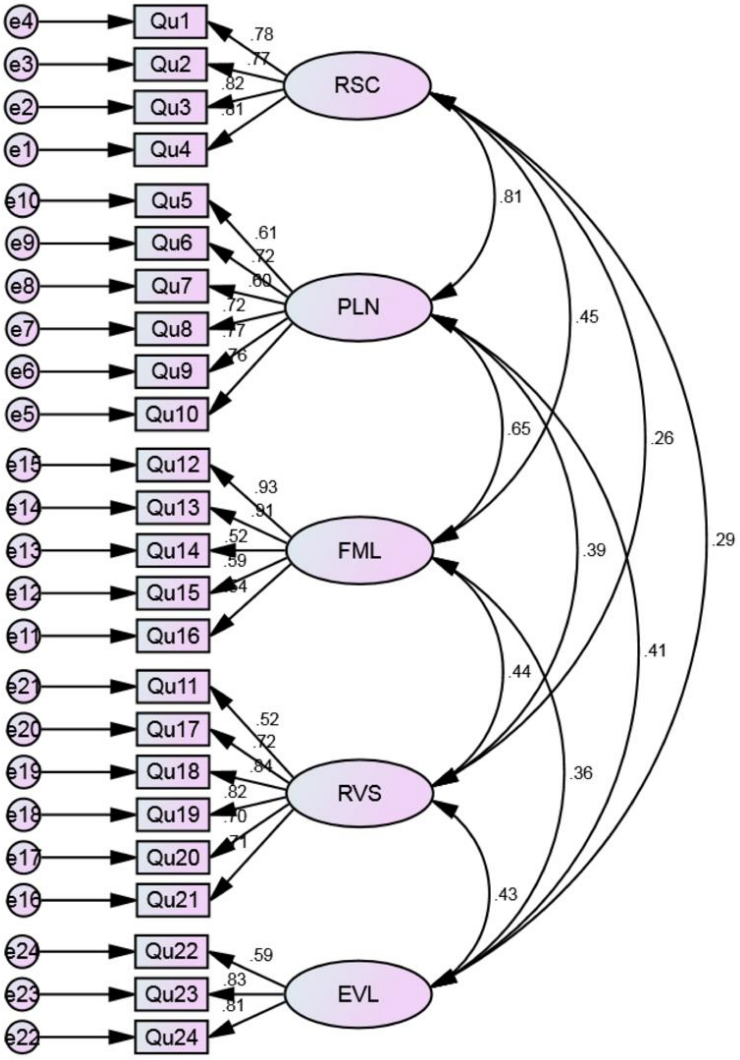
The standardised estimates output of the correlation coefficients between the 24 tested items and five latent variables and those between the five latent variables themselves are shown in Figure 4.1 below. In this figure, the five large ovals represent five identified factors that may define students' composition strategies: research (RSC), planning (PLN), formulation (FML), revision (RVS) and evaluation (EVL). The double-headed arrows between them indicate that the five factors were correlated to each other to different degrees (as shown by the

corresponding correlation coefficient values). The single-headed arrows pointing from the ovals to the rectangles indicate that the items can be explained by their respective factors. All the factor loadings were significant and most were above .60, suggesting that the five factors are well represented by these items (Harrington, 2009).

The varying correlations between the five factors as expected in that the composition processes were perceived to be an integrated whole, and all the factors were hypothesised to be parts of the composition processes. There was a high positive correlation between RSC and PLN (.81), indicating a close interconnection between research and planning. On the contrary, there was a weaker correlation between RSC and RVS/EVL (.26/.29), indicating a potentially weaker interconnection between research and revision/evaluation.

Figure 4.1

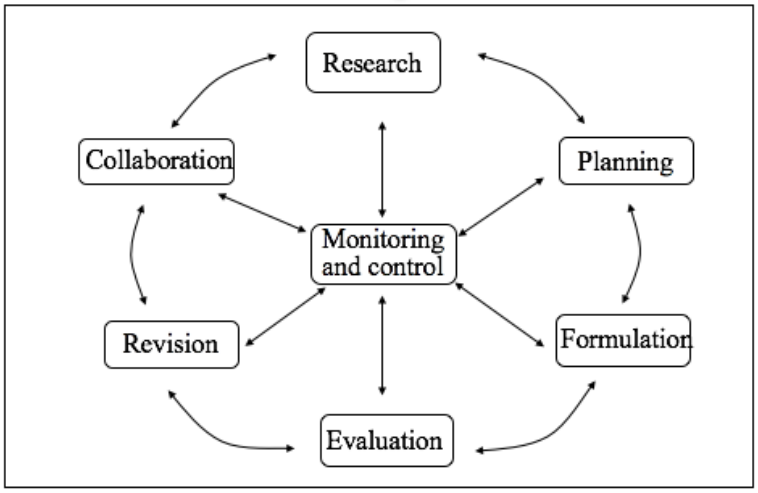
The 24-item, five-factor model for processing strategies



Students' processing strategies were originally conceptualised as consisting of six interactive, inter-related subcategories (as shown in Figure 4.2 below). The results of PCA and CFA from the survey items support five of the subcategories, but not collaboration; the items for this category did not have high correlations.

Figure 4.2

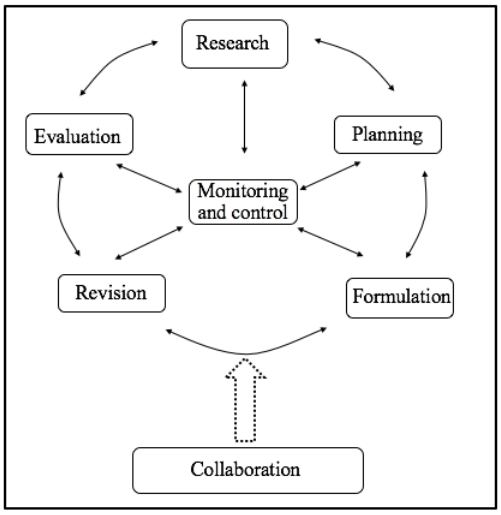
The original processing model



The students' responses to items related to the category of collaboration indicate homogeneity in their collaboration skills. That is, most students tended to cooperate with their peers or seek consultation from peers or teachers when possible, and collaboration may be a shared strength among these students. Based on this result, the conceptualisation of students' processing strategies and the interaction among the subcategories were modified as shown in Figure 4.3. Collaboration was moved outside of the composition cycle, but it still functions as a valuable factor in supporting the whole composition process.

Figure 4.3

The adapted processing model

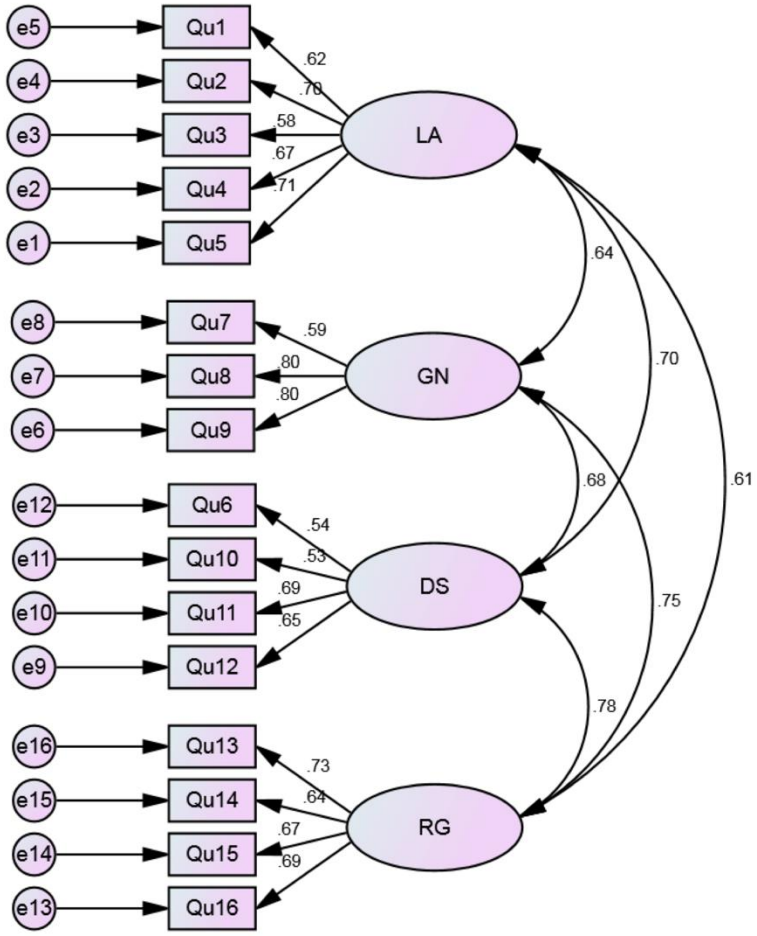


In a similar way, a 16-item, four-factor model was extracted for understanding students' mastery of academic writing knowledge. The results of a forced factor PCA generated four meaningfully interpreted factors with 16 items. Three of the four factors had eigenvalues greater than one, and the disciplinary factor had an eigenvalue of .982. The four factors explained 37.7% (language skills), 9.7% (genre knowledge), 7.0% (disciplinary knowledge) and 6.1% (reasoning and argumentation skills) of the total variance, respectively. In addition, in the four-factor solution, the commonalities of most items were above .5, with three items having eigenvalues of .48 and .49, approximate to .5. The four factors could explain 60.47% of the total variance, meeting both the solution criteria and the interpretability criteria. Moreover, the CFA results from the four-factor knowledge model showed a fairly good model fit ($\chi 2/df = 2.50$, TLI = .88, CFI = .91, RMSEA = .07, SRMR = .06) according to the standards suggested by Browne and Cudeck (1992) and Kline (2015). Cronbach's alpha reached .79 for LA, .76 for GN, .69 for DS and .78 for RG, indicating acceptable reliability. As a result, the 16-item, four-factor model was retained for the analysis of students' academic writing knowledge.

The standardised estimates output of the correlation coefficients between the 16 tested items and four latent variables and those between the four latent variables themselves are shown in Figure 4.4 below. In this figure, the four large ovals represent four identified factors that may define students' academic writing knowledge levels: language skills (LA), genre knowledge (GN), disciplinary knowledge (DS) and reasoning and argumentation (RG). Following the rules of thumb for CFA tests, all these factor loadings are considered fair, and all the 16 test variables significantly load on the expected latent variable (Harrington, 2009).

Figure 4.4

The 16-item, four-factor model for academic writing knowledge



4.4.2 Strengths/weaknesses in essay writing

RQ 4b. *What are the relative strengths and typical difficulties of students in the composition process and in terms of their academic writing knowledge?*

Students' strengths and typical difficulties/problems were explored via both the six-point Likert scale items and open-ended questions. The details of students' processing strategies and academic writing knowledge levels are shown in Table 4.2 and Table 4.3 below and in Table 4.4 in Appendix F.

Table 4.2

Description of the five-factor, 24-item processing strategies

(n = 46 for the high-proficiency group, n = 122 for the mid-proficiency group, n = 109 for the low-proficiency group, and n = 277 for the total across all groups)

Category	High proficiency		Mid proficiency		Low proficiency		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RSC	4.810	.865	4.849	.853	4.927	.768	4.873	.821
PLN	4.615	.830	4.625	.748	4.757	.779	4.676	.774
FML	4.544	.981	4.750	.803	4.695	.767	4.694	.822
RVS	3.986	.980	4.029	1.014	3.880	1.011	3.963	1.006
EVL	4.189	1.174	4.371	1.005	4.159	1.086	4.258	1.068
PO	4.424	.714	4.508	.659	4.480	.624	4.483	.653

RSC: Research, PLA: Planning, FML: Formulation, RVS: Revision, EVL: Evaluation, PO: Processing strategies overall

Table 4.2 above shows the average mean score for each of the five identified processing strategies among participants as a whole and across three proficiency levels. In this table, PO represents the overall processing strategies. The score on an item represents the level of the writer's commitment to this particular composing strategy/activity. The higher the score, the more commitment the student was likely to make to this category. The mean score for PO was 4.483, suggesting that students on average had a basic level in applying the composition strategies. The set of strategies with the highest mean value was RSC (mean = 4.873) and that with the lowest mean value was RVS (mean = 3.963). This suggests that research may be the activity at which the students applied most, whereas revision may be the activity that the students used least. Furthermore, according to the mean values, the student commitment or application among the five processing strategies may be as follows: RSC (mean = 4.873), FML (mean = 4.694), PLN (mean = 4.676), EVL (mean = 4.258) and RVS (mean = 3.963). In addition, the values of standard deviation for revision and evaluation were higher than those for the other three strategies. This suggests that students displayed more variation in their application of revision and evaluation strategies.

Table 4.3 below shows the average mean score for each of the four identified academic writing knowledge components. In this table, KO represents the overall academic writing knowledge. The score on each item represents level of mastery on a particular category. The mean score for KO was 4.186, suggesting that students on average had a basic mastery of academic writing knowledge. The category with the highest mean value was RG (mean = 4.230) and that with the lowest mean value was LA (mean = 4.125). This suggests that students might have better knowledge and skills in reasoning and argumentation but relatively poor language skills.

Table 4.3

Description of the four-factor, 16-item knowledge features

(n = 46 for the low-proficiency group, n = 122 for the high-proficiency group, n = 109 for the low-proficiency group, and n = 277 for the total across all groups)

Category	High proficiency		Mid proficiency		Low proficiency		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LA	4.226	.703	4.160	.708	4.042	.698	4.125	.709
GN	4.145	.923	4.177	.882	4.254	.854	4.202	.876
DS	4.125	.821	4.203	.818	4.236	.734	4.203	.785
RG	4.103	.862	4.236	.828	4.278	.735	4.230	.798
KO	4.156	.670	4.194	.659	4.190	.578	4.186	.628

LA: Language skills, GN: Genre knowledge, DS: Disciplinary knowledge,
RG: Reasoning and argumentation, KO: Academic writing knowledge overall

The results of the quantitative survey items were largely similar to the results of the qualitative data analysis based on the open-ended questions, as shown in Table 4.4 in Appendix F. This table shows the frequency percentage distribution of participants' perceptions of their own processing strategies, academic writing knowledge and other factors

that may affect their writing. As seen in this table, the three most frequently reported strengths were idea organisation in planning (26%), logic in reasoning and argumentation (18%) and research of related sources in research (17%). The three most frequently reported weaknesses were lexical complexity (58%) and grammar (30%) in language skills and disciplinary background knowledge (14%) in disciplinary knowledge. More than half of the students (58%) expressed their need to improve their vocabulary related to their disciplinary studies. These results are consistent with those of the quantitative survey item analysis, as shown in Table 5.3 above.

In addition, only 1% of the students claimed sufficiency of revision and monitoring of the whole composition process as their strengths. Meanwhile, 5% and 3% of students claimed orchestration of processing strategies and monitoring of the whole composition process as their particular problems/difficulties, respectively. That is, around 5% of students did not have a clear understanding of the composition processes and displayed less effective time management skills in the orchestration of typical processing strategies. In addition, in alignment with the results from EFL and CFA, the qualitative analysis showed that collaboration was conceived as a typical strength in students' academic writing processes (5%), and no student perceived it as a problem or difficulty.

4.4.3 Similarities/differences at three levels of English proficiency

RQ 4c. *What are the similarities/differences among individual processing strategies and academic writing knowledge across three levels of English proficiency and writing performance?*

The similarities/differences across three levels of L2 proficiency and writing proficiency were examined both through ANOVA analysis of quantitative survey items (in Tables 4.5 and 4.6) and content analysis of open-ended questions (in Table 4.7, Appendix G).

The results of the one-way ANOVA test showed no statistical significance between students across three levels of English proficiency regarding processing strategies or academic writing knowledge. There was also no significant difference in academic writing knowledge across the three writing levels. Statistically significant differences were only found between students with high-level and low-level writing skills in terms of PO and RVS in processing strategies. The details of the comparison are shown in Tables 4.5 and 4.6 below.

As shown in Table 4.5, HL students on average had higher scores in processing strategies than ML and LL students. There were statistically significant differences between HL and LL students regarding PO (4.606 vs. 4.335, p = .045) and RVS (4.156 vs. 3.656, p = .013), as shown in Table 5.6. This indicates that HL students might apply more revision strategies and overall processing strategies than LL students. Meanwhile, there were roughly similar trends in student processing strategies across the three levels of writing, with RSC being the strongest feature and EVL and RVS being relatively weaker features. That is, students across all three writing levels tended to make more commitment to research, planning and formulation activities than to evaluation and revision activities.

Table 4.5

Description of mean values among the five processing strategies across three writing levels (n = 107 for the high-level group, n = 120 for the mid-level group, n = 46 for the low-level group)

Category	High level		Mid level		Low level	
	Mean	SD	Mean	SD	Mean	SD

PO	**4.606**	.614	4.447	.664	**4.335**	.634
RSC	4.949	.743	4.896	.794	4.717	.991
PLN	4.801	.692	4.622	.808	4.569	.796
FML	4.798	.767	4.627	.879	4.674	.725
RVS	**4.156**	.892	3.923	1.043	**3.656**	1.103
EVL	4.331	1.109	4.240	1.058	4.145	1.017

Table 4.6

Comparison of mean values of processing strategies across three writing levels

(n = 107 for the high-level group, n = 120 for the mid-level group, n = 46 for the low-level group)

Subject	HL vs. ML			HL vs. LL			ML vs. LL		
	t	p	SE	t	p	SE	t	p	SE
PO	.1.86	.151	.085	**2.44**	**.045**	.113	.99	.571	.111
RSC	.52	.877	.108	1.59	.241	.143	1.21	.415	.141
PLN	1.78	.186	..101	1.81	.198	.134	.38	.914	.132
FML	1.56	.252	.108	.93	.661	.143	-.32	.940	.141
RVS	1.80	.185	.133	**2.71**	**.013**	.176	1.45	.273	.173
EVL	.64	..797	.142	.97	.588	.189	.52	.867	.186

A qualitative data analysis revealed a similar trend but with greater detail, as shown in Table 4.7 in Appendix G. According to this table, there were more similarities than differences among students across the three proficiency levels in terms of their self-perception of their strengths and weaknesses, especially concerning processing strategies. For students at all three proficiency levels, lexical complexity (around 60% of the participants) and grammar (≥16% of the participants) tended to be shared problems/difficulties, whereas research of related sources and logic in argumentation (about 20% for both categories) were perceived as shared strengths. In addition, both HP and LP students reported idea organisation as a strength (about 30% for both groups) and a weakness (20% for HP and 14% for LP group), but they perceived idea organisation more often as a strength than as a weakness.

Altogether, consistent findings were obtained from analyses of quantitative survey items and qualitative open-ended questions. Students on average exhibited basic processing strategies and academic writing knowledge. Among all the knowledge categories and processing strategies, language skills were the weakest aspect of students' writing, especially in terms of disciplinary vocabulary and grammar. Within the processing strategies, students reported relative strengths in researching, especially with respect to research of related sources, and relative weaknesses with revision and evaluation, especially in terms of insufficiency of revision, orchestration of the processing strategies and monitoring of the whole composition process. There were no statistically significant differences among students across the three L2 proficiency levels, which indicates that L2 proficiency may not have a direct impact on the use of processing strategies. Alternatively, the classification of students' L2 proficiency may not be accurate. Meanwhile, there were significant differences regarding overall processing and revision between high- and low-level writing, suggesting that more skilled processing strategies and revisions may result in better writing performance.

4.5 Discussion

This chapter has mainly discussed students' writing strengths and weaknesses as revealed by the survey study. First, their strengths and weaknesses were discussed within the confirmed five-factor processing model, then within the confirmed four-factor academic writing knowledge model. Afterwards, similarities, differences and potential interactions were discussed among these two frameworks.

This survey study provides rich empirical evidence in understanding EFL learners' composition processes in academic writing. The processes of EFA and CFA yielded a five-component processing model that could represent students' composition processes. Compared

with the original theoretical model as reviewed earlier, the collaboration strategy fell out of the model. Collaboration as a processing strategy for individual writing has been frequently employed and has been examined in previous studies (de Guerrero & Villamil, 1994; Stapleton, 2010). De Guerrero and Villamil (1994) considered collaborative interaction a beneficial strategy, especially in the form of peer revision to complement independent intellectual functioning. Students in the present study also perceived this strategy as a strength of theirs. Stapleton (2010) suggested having collaboration as an integral part of the composition process, especially in an electronic age. In his case study, the participants spent 8% of the total composition time seeking advice from peers and mentors. Similar to Stapleton's (2010) findings, almost all participants in the present study spent time collaborating with others. In fact, collaboration is so prevalent and has such close interactions with other processing strategies that it could not be distinguished as a separate composition process. Thus, this study proposes a new model to capture the role of collaboration in writing (as shown in Figure 5.3 above). In view of the significance of collaboration, the present study suggests that collaborative revision may be a particularly beneficial strategy in individual writing and may thus be worthy of focused instruction and practice in L2 writing courses.

Within the processing strategies, students' typical weaknesses were in relation to sufficiency of revision, orchestration of the processing strategies and monitoring of the whole composition process. That is, these students did not have a clear understanding of the composition process and displayed less effective time management skills during the orchestration of typical processing strategies. Compared with studies on composition strategies, the time dimension of writing remains under-researched. Existing studies suggest that skilled writers are advantaged over unskilled writers regarding allocation of time to the various processes (Byrnes & Manchon, 2014; Roca de Larios et al., 2008; Stapleton, 2010).

In particular, with increasing L2 proficiency, more balanced allocation of time can be employed (Roca de Larios et al., 2008), especially between the three core activities of planning, formulation and revision. This was partly true in the present study, as seen in the mainland participant Lydia's case in the multiple case study. On the other hand, the multiple case study also showed that Chloe, a local Hong Kong participant, whose L2 proficiency might not be as high as Lydia's, also demonstrated efficient time management skills in the composition process, albeit with less care during the revision process. The secondary school learning experience was likely to be a key contributing factor in that Chloe had a much richer experience with extended writing than her mainland peers before going to university. Thus, besides improvement of L2 proficiency, constant academic writing practice, along with an emphasis on time management skills among the core composition activities, may be critical to students' acquisition of proper revision strategies and monitoring skills.

Furthermore, special awareness and proper allocation of time should be given to revision, especially for low-proficiency L2 EFL learners. Peer revision in particular would be a practical strategy. The present study found that across the three English proficiency levels, research was perceived as the strongest feature and evaluation and revision as weaker features; furthermore, LP students reported greater efforts in research than MP and HP students, but they were more unsatisfied with their evaluation and revision strategies. Content analysis of the open-ended questions shows that one factor contributing to LP students' difficulty with revision is that they could not identify the problems in their own writing, whereas many studies suggest that proficient writers tend to exhibit relatively quick planning, rapid synthesis of ideas or formulation of content and reservation of an appropriate amount of time for evaluation and revision (Baaijien & Galbraith, 2018; MacArthur, 2018). In other words, L2 EFL learners with low language proficiency may have greater need for instruction

in evaluation and revision strategies. Revision strategies in writing have been a constant concern in studies on writing (Chen & Zhang, 2019; Hayes, 2012). Viewing writing as both a problem-solving process and a discovery activity, Baaijien and Galbraith (2018) suggested that greater emphasis should be placed on revising the global structure of the text than on revising sentence forms; this revision focus may provide students with clearer guidance. Meanwhile, it is also necessary to enrich students' knowledge of typical writing problems and effective writing criteria, for which the proposed diagnostic assessment rubric, as discussed in Chapter 2, Study 2, may serve as a helpful tool in facilitating students' self or peer evaluation and revision.

Regarding academic writing knowledge levels, the four-dimensional model can account for students' knowledge levels, language skills (LA), genre knowledge (GN), disciplinary knowledge (DS) and reasoning and argumentation. This four-dimensional model is in line with Snow and Uccelli's (2009) theoretical conceptualisation, whereas Zhao (2019b) found that a two-dimensional model (i.e., content knowledge and language skills) may well represent Chinese undergraduates' perception of academic writing constructs. Similar to Zhao's (2019b) observation, participants in the present study had weak genre awareness and knowledge in academic writing; however, in contrast to the results obtained by Zhao (2019), participants in this survey study showed a relatively clear understanding of the key components of academic writing, especially regarding reasoning and argumentation. One contributing factor may be the writing context. In this study, the survey was conducted primarily within a business study disciplinary learning context, whereas the majority of participants in Zhao's (2019b) study were majoring in English. In other words, students from different disciplinary backgrounds are likely to hold different perceptions of academic writing knowledge concepts, especially in language courses and in disciplinary courses. Thus, it is

necessary to facilitate the transfer of students' genre knowledge across courses. Proper adoption of genre-based approaches from the ESP perspective may have special advantages in scaffolding the development of such transferring abilities.

In addition, in the survey study, lexical complexity was identified as the most desired quality of writing. More than half of the students (58%) expressed their need for a larger vocabulary related to their disciplinary studies. In many studies on L2 writing, lexical complexity has been identified as a marker of effective writing by secondary school and college students (Crossley, 2020; Crossley et al., 2014; MacArthur et al., 2019) and even as the single most important area of L2 competence in academic achievement (Hyland, 1997; Saville-Troike, 1984). Meanwhile, discipline-related vocabulary has been ranked as the most problematic aspect of academic English by L2 EFL undergraduates (Hyland, 1997; Evans & Green, 2007). Thus, it is proposed that a discipline-specific common core lexis should be given special emphasis in EAP programs. One important benefit of lexical complexity, particularly for L2 learners, may be that this level of language proficiency may reduce students' cognitive workload and free them to work on content and arguments. Hence, it is necessary to focus on supporting student improvements in lexical complexity in the development of DC in L2 EFL writing courses. Considering the developmental factors in vocabulary acquisition, one typical suggestion is to foster student autonomy in learning vocabulary (Evans & Green, 2007) and another good method is for institutions to provide proper language support throughout students' university studies, especially in L2 learning contexts (Hyland, 1997). Specialised EAP courses may have the capacity to provide language support and to help students develop discipline-pertinent writing features (Rose et al., 2019).

This study also identified a strong positive correlation between research and planning and a weak correlation between research and revision/evaluation, with planning also having a high correlation with formulation. This may suggest that research and planning strategies have a salient influence on the formulation process, and effective research and planning activities can greatly facilitate the formulation process. Stapleton (2010) specified that research is an integral part of the idea-generation process, which connects directly to the planning and formulation of a written product. That is, the core concern of idea generation may largely account for the high correlations between research, planning and formulation. Thus, it is necessary to focus on instruction in research and planning strategies in L2 EFL writing courses to facilitate a smooth formulation process and reserve sufficient time for later revision and evaluation. Idea generation may serve as a meaningful objective in connecting research, planning and formulation strategies. In this respect, disciplinary context would be an effective scaffold for the integration and iterative exercise of these three strategies.

Last but not the least, it was noted that no statistically significant difference was found in the perception of most of the process strategies (with the exception of overall processing strategies and revision strategies) and all the academic writing knowledge dimensions, between students of different English or writing proficiency levels. On one hand, it may indicate the relative independence of some strategy use from L2 proficiency in L2 writing or the great homogeneity of novice L2 learners' perception on academic writing knowledge, as observed by Zhao (2019). In other words, the novice L2 student writers may not have a more fine-grained understanding of the composing process and academic writing constructs. On the other hand, this weak differentiation power may raise questions about the validity of the constructs, and about the affordance of questionnaire as a data collection instrument for self-reporting. Survey and questionnaire instruments have been frequently employed in L2 writing

studies for examining learners' writing strategies (e.g., Qin & Zhang, 2019; Teng & Zhang, 2016) or knowledge status (Zhao, 2019). Some of these survey studies only briefly mentioned the predictive power of the identified factors on writing efficiency (e.g., Teng & Zhang, 2016), but did not delineate clearly the differentiating power between high, medium-, and low-proficiency writing. Still another possibility is that students have insufficient self-evaluation skills. Thus, consideration of writing proficiency difference and cultivation of L2 learners' self-evaluation skills are two issues worth further consideration.

4.6 Conclusion

The present study identified five factors representing students' composition processes (research, planning, formulation, revision and evaluation) and four factors accounting for students' academic writing knowledge (language skills, genre knowledge, disciplinary knowledge and reasoning and argumentation). Students on average displayed a basic level of mastery of processing strategies and academic writing knowledge, and there were more similarities than differences between students across the three proficiency levels, especially in terms of processing strategies. Subject-related vocabulary was identified as the most desired quality in student writing. Meanwhile, collaboration was a helpful and more commonly self-perceived strength, but it was not considered essential in students' academic writing process. In addition, there were significant differences between high- and low-level writing in terms of overall process and revision. In view of these difficulties and underlying causes, the provision of a specialised EAP course, along with proper implementation of the process-genre-based approaches from the ESP perspective, may be critical to students' acquisition of proper composition strategies and enhancement of students' academic writing knowledge.

This study has several limitations. The survey was conducted shortly after the participants had performed a 500-word discipline-specific writing practice. Thus, the students' perception of their composition strategies and academic writing knowledge was more likely to be confined to this specific disciplinary essay writing context. It may be meaningful to apply the survey items to a slightly different context—for instance, to different lengths of essay writing or to essay writing in a different course—to perceive the potential similarities and/or differences across contexts. In addition, the targeted student population is relatively restricted, and the sample size of the survey is limited. Thus, the interpretation of the results should be considered with caution.

The survey study in this chapter reveals students' processing strategies and academic knowledge levels on a larger scale. The previous case studies in Chapters 3 and 4 provide a richer, more nuanced understanding of students' composition processes. The textual analysis in Chapter 2 identifies specific discourse-level problems. Based on the diagnosis of students' discourse problems and potential problems in the writing process, an intervention study was conducted to improve students' writing. The next chapter reports on this intervention study.

Chapter 5 Study Four — An Intervention Study

Chapter Abstract

This intervention study focused on a metacognitive approach for improving EFL learners' global coherence in essay writing, especially their performance in the introduction and conclusion sections. Two intact classes majoring in business studies participated in the study, one as the experimental group and the other as the control group. A total of seven hours' intervention was implemented over a 16-week reading and writing course, including one orientation session and three iterations of a set of metacognitive activities. One-way ANOVA analysis found some significant improvement by the experimental group in the post-test; however, the control group slightly outperformed the experimental group in the delayed post-test. Potential factors contributing to the unexpected findings were explored via qualitative interviews. The present intervention process further reinforced the necessity for teachers/trainers of academic writing to have adequate disciplinary backgrounds so that they can deliver effective academic writing instruction and facilitate undergraduates' transferring of related knowledge and skills across courses.

5.1 Introduction

Informed by previous case studies and survey studies, the present chapter will report findings from an intervention study to address one of the typical discourse issues. Considering the importance of sequencing in instruction and the significance of global coherence in academic writing, as discussed in chapter 2, this intervention study focuses specifically on the development of global coherence (GC) in writing among the five main discourse categories. GC in writing is an essential aspect of discourse competence (DC). GC pertains to the overall conceptual structure of a piece of writing and carries the rhetorical pattern characteristic of and appropriate to the communicative purpose (Bachman, 1990; Lee, 2002). Owing to its

importance to good writing, GC has been researched extensively within different theoretical frameworks, such as macrostructure (Lee, 2002), discourse organisation (Hoey & Winter, 1983) and genre analysis (Swales, 1990).

This intervention study was conducted to explore the effectiveness of purposefully designed metacognitive instruction on the GC of a group of Chinese EFL undergraduate students, focusing on the knowledge and skills that they exhibited in writing the introduction and conclusion parts of their academic essays. Altogether, three iterations of metacognitive instruction were delivered over 16 weeks with the experimental class. The following section will provide a summary of the review of existing studies on coherence and metacognition instruction which provides the conceptual foundation for this study.

5.2 Literature review

This section first reviews approaches to teaching global coherence in L2 writing. It then focuses on the application of metacognitive instruction in L2 learning and reviews existing intervention studies. Based on the synthesis of typical approaches in teaching global coherence and methods of metacognitive instruction, the study designed a set of metacognitive instruction procedures, implemented the design and evaluated its effectiveness.

5.2.1 Teaching of global coherence in L2 writing

The primary concern of GC is organisational patterns, a concept much like that of moves in genre studies (Hyon, 2016). Hence, teaching GC is closely related to the genre-based approach in teaching and learning academic writing. Three approaches to genre studies can be identified in the literature: Australian systemic functional linguistics (the SFL perspective), American New Rhetoric studies, and the English for specific purposes (ESP) perspective (Hyland, 2004; Hyon, 1996; Swales, 1990; Tardy, 2009).

The three approaches had different educational contexts, intellectual goals and instructional frameworks. In general, the SFL approach emphasises genre as ‘a staged, goal-oriented social process’ (Rose & Martin, 2012) which is intended to create a genre-based pedagogy consistent with the theoretical work of SFL theories. The New Rhetoric approach has as its main concern L1 teaching and emphasises genre as situated, dynamic, flexible, free, open to changes and subject to negotiation. Consequently, this approach draws on social and literary theories instead of linguistic theories and prefers ethnographic methods to the study of genres. The approach may provide language teachers with a richer view of the social, cultural and institutional contexts of the genres. The ESP approach emphasises genre as professional competence and focuses on the communicative needs of non-native English speakers in academic and professional contexts. The ESP approach sees genre as a class of structured communicative events characterised by specific communicative purposes and various patterns of ‘structure, style, content and intended audience’ (Swales, 1990, p. 58). That is, it is concerned with both the social functions and the linguistic and rhetorical forms of language use.

In the strand of literature taking the ESP perspective, the seminal work is Swales’s (1990) genre-based approach to understanding, teaching and learning English in academic and research settings. In his work, a genre is understood as ‘a class of communicative events, the members of which share some set of communicative purposes’ (Swales, 1990, p. 58). And ‘purpose’, rather than similarities of form or some other criterion, is considered the primary determinant of genre membership. Specifically relating to research work for undergraduate students, Nesi and Gardner (2012) proposed five main purposes of academic writing in university study: (1) demonstrating knowledge and understanding, (2) developing powers of

informed and independent reasoning, (3) developing research skills, (4) preparing for professional practice, and (5) writing for oneself and others. They analysed 2,761 assignments written by university students across disciplines and identified 13 genre families: case study, critique, design specification, empathy writing, essay, exercise, explanation, literature survey, methodology recount, narrative recount, problem question, proposal and research report. Each genre family serves a unique purpose and has its own structure. These 13 genre families further comprise more than 100 specific genres. This systemic and comprehensive classification provides practical guidance for the development of teaching and learning materials, assessment design and instructional strategies in English-medium instruction university study.

Another influential figure in genre studies in L2 writing is Hyon (1996; 2016), who introduced six activities to raise students' genre awareness and help them acquire specific genre types in a discipline-specific writing course (Hyon, 2016). Her first activity was to introduce genre and moves through a dance lesson; the second activity was to identify moves in a specific genre type (e.g., book reviews). The third activity aimed to develop a repertoire of evaluative vocabulary and disciplinary values concerning a particular genre type. The fourth was to interview disciplinary experts to understand the context within which a particular genre is adopted. The fifth was individual writing practice. The sixth was a comparative analysis of a different genre type in which students are required to highlight the differences between the two genre types. The six activities were appropriate for undergraduate EFL learners who were new to academic writing and disciplinary genre types. They provided the present research with a valuable reference for the design of the intervention study.

Johns's works (1997; 2008; 2011) were also of great importance. Given the elusive and dynamic nature of genre, Johns (1997) suggested that genre teaching should aim not at correct reproduction of certain conventions but at providing methods and techniques for genre analysis. Subsequently, the teaching should begin with textual features (as the starting point and explicit focus) but move rapidly towards integration of theories and practices (such as genre learning) that value the analysis of context, complex writing processes, and intertextuality (p. 64). Such analysis could result in high transfer. Furthermore, for genre analysis, Johns suggested thorough analysis followed by critique or evaluation activities (Johns, 1997; 2011). Specifically, Johns (1997) proposed three main goals in teaching genre: (1) applying genre knowledge for analysis and criticising known and new texts, (2) evaluating, expanding and revising strategies for approaching genre, and (3) developing a set of meta-languages for genre analysis. For teaching EFL learners in university study, she suggested beginning with daily life L1 genres that students were familiar with, then moving to comparable L2 genres, and afterwards proceeding to genres they needed to master in L2 writing. She encouraged students to explore genres of their own interests and collect genre exemplars. Johns's approaches provided valuable techniques for supporting autonomous learning.

More recently, Tardy et al. (2020) provided a systemic reference on the principles for genre-based pedagogies. Their study proposed a comprehensive theoretical framework from a social-cognitive perspective and articulated five key constructs of genre knowledge and their interrelations: genre-specific knowledge, genre awareness, metacognition, recontextualisation and multilingualism. Genre-specific knowledge and genre awareness are both permeable to and distinguished from each other; together, they constitute a broader construct of genre knowledge. Genre awareness is also part of metacognition. Effective use of genre could take

the form of the adaptation of existing genre knowledge to new or unfamiliar writing situations, that is, recontextualisation. All these activities for genre use should be situated in a specific social context, which often involves two or more languages, that is, a multilingual setting. Elaborated differentiation between and proper integration of these five key constructs helps develop students' rhetorical flexibility within and across courses and languages.

Besides the specific exploration of genre-based approaches reviewed above, there were other proposed activities. Hyland (2004) synthesised some typical genre-based approaches to writing, including scaffolding (with reduced teacher support), collaboration (peer interaction), the teaching-learning cycle, and consciousness-raising. The teaching and learning cycle can comprise five main stages (Feez, 1999; Hyland, 2004): developing the context, modelling in deconstructing the text, collaborative construction of the text, independent construction of the text, and linking related texts beyond. These five stages are intended to be used flexibly, with free selection of certain stages for learners of different levels or with potential revision of certain stages.

Despite the abundant and substantial research outcome, emphasis on the genre-based approach alone may undervalue processing skills and other related textual or contextual components of writing. To better facilitate student development of writing abilities, diverse perspectives must be incorporated into classroom teaching. For instance, instruction on processing strategies and learners' needs should be integrated into the multilingual setting. Also, existing research on genre-based instruction focuses on L2/EFL learners studying in EMI universities overseas and English-as-L1 university students. It is necessary to examine the exercise of EAP writing by L2/EFL learners in a local non-English culture.

5.2.2 Metacognitive instruction in language learning

Metacognitive instruction aims at enhancing metacognition in practical teaching and learning (Sato & Leowen, 2018). Metacognition comprises three major components: metacognitive knowledge, metacognitive strategies, and metacognitive experience (Flavel, 1976). Hacker et al. (2009) considered writing as a field of applied metacognition. In view of writing from this perspective, the present study proposed that instruction in writing may be conceived in three dimensions: writing knowledge, writing strategies and writing experience. Metacognitive instruction was found effective in improving L2 listening (Cross, 2011), reading (Zhang, 2010) and speaking (Zhang et al., 2021), as well as in content area studies (Hacker et al., 2019; Jumaat & Tasir, 2016). However, metacognitive instruction has scarcely been explored explicitly in L2 writing research (Ruan, 2014; Sato & Loewen, 2018).

Existing studies from related perspectives have provided rich references for practical procedures in implementing metacognitive instruction (Hacker et al., 2019; Sato & Loewen, 2018; Teng & Zhang, 2020). Sato and Leowen (2018) investigated the effects of metacognitive instruction on feedback with a group of university-level EFL learners in Chile. Their metacognitive approach consisted of five stages: (1) introducing the definition and objectives of corrective feedback, (2) explaining the effectiveness of corrective feedback concerning relevant theories, (3) class discussion (addressing student questions), (4) looking out for feedback, and (5) reminding the students of corrective feedback at the end of the intervention. Their results showed that metacognitive instruction could enhance the effects of corrective feedback in promoting L2 development. Hacker et al. (2019) designed a language-based metacognitive instruction intervention to help upper elementary students with mathematical reasoning using mathematic language while promoting self-regulated learning. Their approach was based on self-regulated learning and was operationalised into six steps: *activating* background knowledge, *discussing*, *modelling*, *memorising*, *scaffolding* and

individual *practising*. Evidence from their study showed that this approach could help students become self-regulated learners who are proficient. They suggested that metacognitive components can be best learnt when embedded into a specific domain and instructed systemically. Teng and Zhang (2020) conducted a quasi-experimental study among a group of EFL undergraduates in China to learn the effect of self-regulation on writing efficiency. They adopted the same procedure as practised in Hacker et al. (2019). Their metacognitive intervention was found effective in increasing self-efficacy in learning and promoting positive academic outcomes.

An important commonality of these studies is that they all engaged in five typical procedures: introduction of the content topic, group discussion, collaborative practice, individual practice, and reflection/memorisation. These common procedures might be primary in effective metacognitive instruction.

Apart from the demonstration of specific intervention stages, the systemic conceptualisation of specific metacognitive strategies and operational sequencing of strategy instruction also shed valuable insights on the implementation of metacognitive instruction. Zhang et al. (2021) explored a group of 136 Chinese EFL undergraduates' use of metacognitive strategies in completing a series of integrated speaking tasks via self-report inventories, self-rating scales and semi-structured interviews. They proposed using a four-dimensional metacognitive strategy instruction involving planning, problem-solving (as the first step and the centre), monitoring and evaluating in an EFL speaking class. They emphasised the importance of sequencing in teaching strategies in particular, beginning with a single metacognitive strategy (with problem-solving strategy as the first) before moving on to teaching and practice on interrelation and interaction among the strategies. Notably, this narrowing-down procedure

for teaching strategies is considered the most effective for facilitating student learning (Plonsky, 2011; 2019).

Some studies proposed a series of principles for the use of metacognitive instruction. Jumaat and Tasir (2016) developed a framework of metacognitive scaffolding (MS) for assisting online learning in a subject course. Through repeated multiple case studies, they identified five mechanisms of MS: focusing on the process of learning (as the dominant principle), followed by presenting a rationale for tasks and activities, supporting reflective writing, encouraging relationships among participants, and supervising text comprehension. They suggested that these five mechanisms can serve as core teaching strategies for maximising students' learning online.

In addition to experimental or quasi-experimental studies, there were also qualitative classroom inquiries. Lee (2002) conducted a case study by teaching coherence to 16 first-year university students in Hong Kong. She conceived coherence as a set of coherence-creating devices (including macrostructure for global coherence) and suggested that teaching the macro-elements of writing should precede teaching the micro-elements. Her study employed a four-episode instructional procedure for each of the identified coherence features: topic introduction by discussion and explanation, consolidation of topic knowledge through the reading of handouts, mini-text analysis for raising coherence awareness, and follow-up (re)writing practice. This approach was found to have broadened the students' understanding of writing and equipped them with diverse resources to facilitate their writing. However, Lee's study conceived coherence as a set of generalised principles, and it was meaningful to explore students' application of coherence features in specific text types or genres.

Another interesting study is that of Yu et al. (2020), who conducted a large-scale survey to understand the impact of three typical writing instruction approaches on students' motivation and engagement. They found that the product-oriented approach induced both adaptive motivation (especially for exam-oriented EFL writers) and maladaptive motivation (i.e., negative feelings), as well as both more and less engagement; the process-oriented approach led to student engagement in writing (with multiple reviewing and drafting opportunities), but with limited impact on motivation (due to the heavy writing load and feedback issues). Finally, the genre-oriented approach could promote both students' adaptive motivation and their engagement in L2 writing (with reduced anxiety and enhanced confidence by the provision of explicit structure). As a result, they proposed a synthesis of various forms of approaches in L2 writing instruction, using a genre-based approach as supplementary.

This literature review leads us to conclude that the metacognitive instruction approach could be useful for enhancing L2 writing performance. Also, systemic synthesis of these approaches may enhance learner motivation and engagement in L2 writing. This is the approach to be adopted by the present research in the design of the intervention study.

5.3 Research design and methods

This intervention study addressed RQ5 of the thesis project: Is the metacognitive instruction approach effective in facilitating EFL undergraduates' development of discourse competence in the classroom setting? Three sub-questions were designed to guide this intervention study:

RQ5a: Does the quality of students' global coherence improve after metacognitive instruction, especially regarding their writing of introductions and conclusions?

RQ5b: Are there any differences in global coherence between students who received the

metacognitive instruction and those who did not?

RQ5c: What are the factors affecting the intervention effect?

5.3.1 The teaching and learning context

This intervention study was conducted in an international joint education programme at a comprehensive university in mainland China. This transnational education programme has 20 years' experience in cooperation, and it has recently launched a four-year full cooperation project with an overseas university. Students enrolled in the four-year programme can choose to complete their bachelor's study at the host university in mainland China or have six months or more of study experience at the partner university overseas. The programme has an enrolment of about 500 students each year, and currently it provides three majors: international business and trade, accounting, and computer engineering. At the time of this study, the programme had approximately 2000 students enrolled.

All students need to fulfil both international and national curriculum requirements, achieving sufficient credits in language courses, bilingual disciplinary courses, courses concerning the law, moral education and politics, and other compulsory courses. The business course in this joint program, i.e., *Integrated Business Challenges*, aimed at improving students' intercultural communication knowledge and skills. At the semester of current research period, this course mainly assesses students through four writing activities: two 500-word journal reflection tasks, a business report by team work, and a 500-word individual essay. Given this present research interest in essay tasks by individual work, the 500-word individual essay task was selected as the target intervention task. For this individual essay, students were required to analyse a business scenario using a series of problem-solving techniques and an about-two-page reading materials were provided to inform the specific business case. Their subject teacher would evaluate the individual essay from aspects: application of problem-

solving tools, logical argumentation, referencing, grammar, and paragraph structure. The reading-based requirement of this 500-word essay is in some degree similar to the reading-integrated writing task in TOEFL (Test of English as a Foreign Language).

The present intervention was situated in a comprehensive English Reading and Writing course. A more appropriate site, an English for specific purpose reading and writing course, was not chosen for this intervention study mainly due to two reasons. One was that on the current research site there was only English for general academic purpose reading and writing courses offered to students. Even if there were occasionally titled instructions related to specific business discipline, that mostly only touched on business-related topics on surface while the whole instruction was still language and grammar based. The other reason was that in the joint program the titled business writing course was managed by overseas' patterners and modification of the course was beyond the researcher's control. This English reading and writing course is provided in the first two years, with a total of four hours weekly of in-class learning focused on improving students' reading and writing skills. This course aims at preparing the students for the end-of-course test required by the partner university, the national college English language test expected in China's college English curriculum, the International English Language Test System (IELTS) for some students and help with students' language needs as required in the disciplinary study.

5.3.2 The focal participants

This intervention was conducted among a group of Year 2 business students in their third semester of the reading and writing course. At the time of the study, there were 18 groups, 12 majoring in business and 6 majoring in computers; each group had around 30 students. There were 16 teaching weeks in this semester, and the reading and writing course was delivered

twice a week, each session lasting about 1.5 hours.

A sample of two intact classes majoring in international business and trade were selected as the participants in the study. The experimental group and the control group were assigned by convenience because of class schedules. That is, the first group, which had the first teaching session, was naturally taken as the experimental group, and the group that followed was taken as the control group. In this way, the instructor reserved more consciousness to observe that the intervention details were conducted only with the experimental group and did not appear in the teaching session for the control group. During the 12-week intervention, the experimental group ($N = 31$) received a series of metacognitive instructions (one session of foundation lecture and three iterations of practice) based on the framework proposed by Hacker et al. (2019) and Lee (2002). In contrast, the control group ($N = 30$) received a comparable amount of reading/writing instruction as conventionally practised in the target education programme.

5.3.3 Instructional materials and instruction procedures

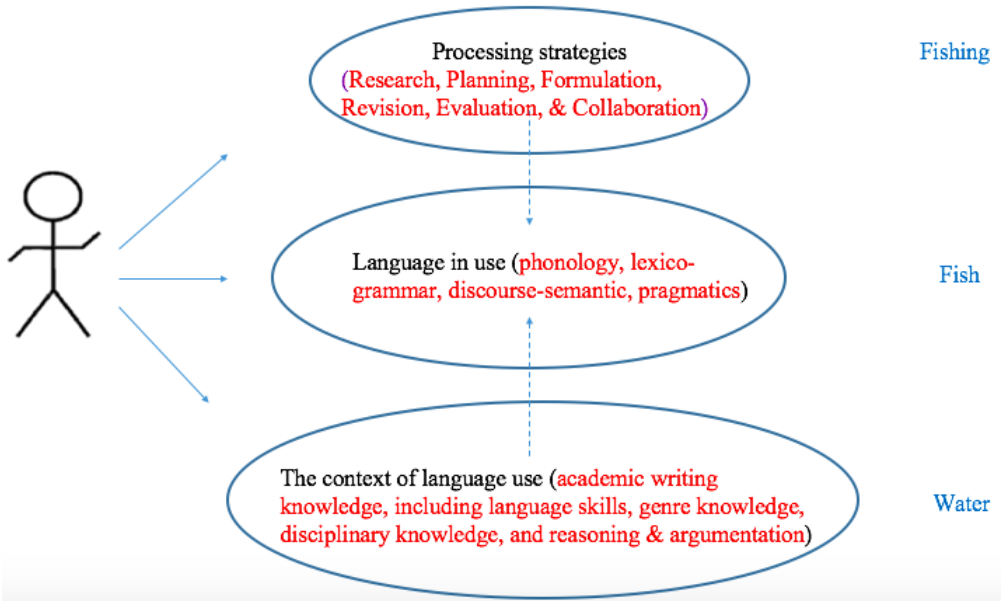
In the English reading and writing course, the routine teaching and learning activities were topic based, and a new topic was introduced every two weeks. During the intervention study, five topics were covered: heroes, society, childhood, technology and arts. An online teaching and learning platform, Ketangpai, was employed for the distribution of related reading materials and the collection of class assignments. The teaching and learning materials for the intervention study were selected based on five criteria: (1) accuracy (verifiable information), (2) authority, (3) objectivity (handling alternative views and interpretations and bias), (4) currency (published in the latest five years), and (5) coverage and scope (related to business and economics studies). Searching for related materials was challenging, as it is difficult to identify proper teaching and learning materials about these five topics (with adequate genre

types and lengths) and fulfil all five selection criteria at the same time, especially regarding the currency of the source. Hence, currency was sometimes not strictly followed, and the length and language of some materials were edited to meet the learning needs.

The present metacognitive instruction approach focused on the integration of three levels of knowledge, skills and practice: academic writing knowledge, processing skills and demonstration of GC in actual language use. Their overall interrelation and interaction are depicted in Figure 5.1 below.

Figure 5.1

The interaction between linguistic features, processing strategies and the academic writing knowledge



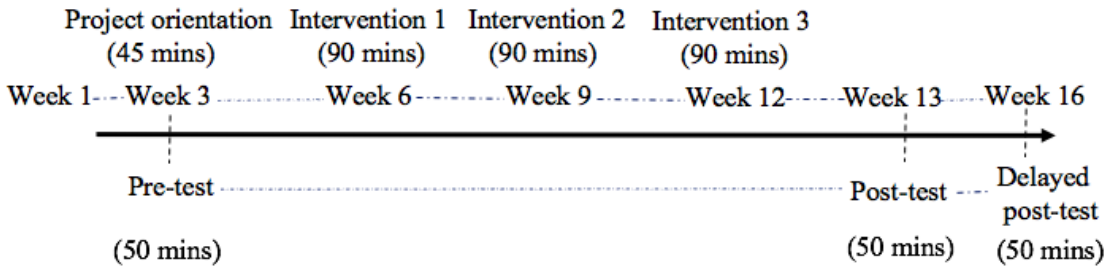
Specifically, the relationship between academic writing knowledge and DC is seen as that between water and fish; the relationship between processing strategies and discourse competence is viewed as that between fishing and fish. To enable a solid mastery of knowledge and skills of discourse competence, students must be equipped with adequate processing strategies and academic writing knowledge.

Specifically, the intervention focused on researching and planning skills for composing and writing the introduction and conclusion sections in an academic essay. The instruction consisted of six moves: *activating*, *discussing*, *modelling*, *memorising*, *supporting*, and *individual practicing*. *Activating* was intended to introduce the target essay type within wider genre families. *Discussing* engaged genre analysis, including analysis of texts, context, and purpose. *Modelling* involved teacher and peer modelling of related processing skills on a specific task. *Memorising* engaged peer work or individual reflection on related genre knowledge and processing skills and reading teacher-prepared handouts. *Supporting* included group interactions and reflections. Finally, *individual practice* was arranged.

At the beginning of the intervention, one orientation session was provided, including an introduction and explanation of a systemic picture of academic writing knowledge and strategies. Based on the preliminary orientation, a sequence of text-context-processing instruction strategies was employed: starting with textual features, focusing on context knowledge and emphasising the iterative practice of related processing skills. The metacognition intervention procedure was iterated three times with the experimental group (EG) throughout the whole semester, while the control class had strengthened teaching and learning activities based on the course conventions. A map of the timelines was drawn for the conception of the intervention activities and the pre-, post- and delayed post-test, as shown in Figure 5.2 below.

Figure 5.2

The timeline of the intervention activities



A sample of the intervention activities is provided in Table 5.1 below.

Table 5.1

A set of intervention activities for learning the Topic on Technology

Objectives	Practice of research/planning skills for introduction/conclusion writing		
Procedures	**Details**	**Purposes**	**Duration of time**
Activation	Read the main body parts of a magazine article and give a title for it (type the proposed title in the online learning platform Ketangpai) and explain how and why the title is related to the provided text	to reinforce the key function and features of introduction/conclusion	10 mins
Discussing	Textual analysis of three styles of introduction/conclusion on the same topic (that in a newspaper report, a book chapter, and a magazine article)	to raise genre awareness by comparison of text, context and purposes in three different genres	20 mins
Modelling	Teacher presentation on rhetorical moves and research/planning strategies in writing an introduction/conclusion for a textbook passage People are from earth Machines are from Outer Space	to scaffold a systemic integration of the textual features and processing strategies	15 mins
Memorising	The key features of introduction/conclusion and activities in research/planning	to facilitate acquisition of discourse knowledge and processing skills	10 mins
Supporting	Group work to explore how and why the introduction/conclusion in the textbook essay were related to the target rhetorical moves	to scaffold peer support for self-regulation in producing effective discourse features	15 mins
Individual Practicing	Add one paragraph before and one after for the given written text that would be published in a business magazine	to develop self-regulation and independence in writing	20 mins

5.3.4 Pre-, post- and delayed post-test materials

Three sets of testing materials were designed and developed for the pre-test, post-test, and delayed post-test, as shown in Appendix H. These three tasks were reading-to-write tasks adapted based on TOEFL bit format writing and IELTS writing. The specific topic or focus and reading materials were developed following three principles: authenticity, effectiveness, and appropriateness. Authenticity means that the task is meaningful to students and relevant

to their experience, interests, and goals. Effectiveness means that the writing task shall facilitate the realisation of course goals and instructional objectives. Appropriateness refers to difficulty, especially regarding topic complexity and lexical complexity. Of these three criteria, one particular challenge is effectiveness of the task design. Despite the researcher's overarching efforts in integrating the intervention in a meaningful way into the students' actual academic writing practices, the three writing topics may still seem somewhat like generic EAP type of writing tasks, instead of the tasks that facilitate students' understanding and application of specific discipline-related concepts. Following these three principles, the writing tasks were designed around three topics: *university students' selection of rooming alone* (for the pre-test), *the conflicts of handwriting and typing* (for the post-test) and *selection of an academic major in university study* (for the delayed post-test). The reading materials were adapted from one book chapter and two academic articles, with slight modifications to some sentences and vocabulary to improve readability.

An online text analysis tool, the Coh-Metrix Common Core Text Ease and Readability Assessor (TERA), was used to examine the readability of the adapted reading materials. TERA assesses the readability of a text from five aspects: narrativity, syntactic complexity, word concreteness, referential cohesion, and deep cohesion (McNamara et al., 2013). A low narrativity score means the text contains more uncommon words and more information and ideas. A low concreteness score indicates fewer concrete and more abstract words; more abstract words might mean more abstract ideas as well. Low deep cohesion scores might indicate a greater challenge to the reader's comprehension of how the ideas, events or information of the text as a whole fit together. At the end of the TERA analysis, an overall reading easibility level was provided on an estimate of a passage grade level using the Flesch-Kincaid grade level readability formula; the higher the level, the more difficult it is to

read. Automatic analysis of the first run of the readability assessment showed that these three reading materials were all average in syntactic simplicity and the amount of deep cohesion. However, they differed slightly in terms of narrativity, word concreteness, and referential cohesion. Referential cohesion in pre-test materials and word concreteness in post-test materials were further modified to improve readability.

Details of the second round of TERA analysis are shown in Table 5.2 below. According to this table, the referential cohesion for the pre-test materials was still low, despite further modification. It was suspected that the narrativity of the materials may be inversely related to the score for referential cohesion, as was shown from the case of analysis for the post-test material. Given that students had been exposed some related reading materials on the same topic, the modified versions were retained for use; as a result, the delayed post-test reading material was slightly more difficult than the pre-test and post-test materials.

Table 5.2

Percentile profile of text easibility

Category	Pre-test material	Post-test material	Delayed Post-test material
Narrativity	47% (average)	34% (average)	13% (low)
Syntactic complexity	60% (average)	63% (average)	66% (average)
Word concreteness	44% (average)	10% (low)	41%(average)
Referential cohesion	13% (low)	19% (low)	50% (average)
Deep cohesion	74% (average)	61% (average)	64% (average)
Flesch-Kincaid Grade Level	10	12	14

5.3.5 Data collection and analysis

Altogether, three types of data were collected for the intervention study: the participants' written essays, the interview protocol, and the researcher's notes on daily teaching and learning experience. RQ5a and RQ5b were addressed by analysis of the written essays; RQ5c were addressed by analysis of the interview protocol and the researcher's notes.

The written essays were collected at three time points: the pre-, post- and delayed post-tests. The pre-test was administered at the beginning of the semester. Both groups received the same writing task accompanied by a reading text. They were required to finish the task in class. After three rounds of intervention activities, the post-test was conducted in a similar situation. Finally, at the end of the semester, the delayed post-test was conducted.

Individual interviews were conducted with three students (Lei, Yu and Xi) from the EG throughout the intervention process. Most interviews were conducted immediately after class intervention and audio recorded. Related notes were taken down, and observations of in-class performance and communication on teaching and learning affairs with other teachers were also documented using a mobile phone app, Diaro.

Participants' handwritten essays were typed into the computer using the Shimo platform (an online document editing and processing tool). The writer's information was removed from each essay and assigned a numeric code to avoid the possible influence of handwriting and writer information on the scoring process. During the transcription process, students' essays were kept the same as the original (including essay organisation, syntax features, spelling, and punctuation).

Two raters—one the researcher and the other an experienced writing teacher—evaluated the essays within two weeks using the discourse rubric validated in Chapter 2. Specifically, given the close interconnection between *topic identification* and *global coherence* and the interaction within the three subcategories of *global coherence*, the first six subcategories of the discourse rubric were applied to evaluate the students' written essays collected across the intervention processes. Before formal scoring, nine essays from the three tasks were selected for trial assessment. The two raters discussed the source texts and writing prompts and then assessed the sample essays together. They reached an agreement on some typical features for differentiating certain discourse levels. They continued to have face-to-face evaluations of the remaining essays over three sessions within one week, with each session focusing on one writing prompt. After the whole scoring process, two sets of scores were cross-checked, and further discussions and decisions were made regarding discourse scores with more-than-one-point differences. Computing of Spearman's correlation showed that there is a strong positive correlation between the two raters' ratings over six constructs (r = .716 for [F1], r = .622 for [F2], r = .848 for [F3], r = .787 for [F4], r = .837 for [F5], and r = .740 for [F6]). The final score for each construct in each essay was the average of the two raters' scores. Also, a discourse overall (DO) category was attained by computing the average of the six sub-discourse features to perceive the overall discourse tendency.

The hypothesis was that the EG's performance would increase over time while the trend in the control group (CG) would be less impressive. Thus, one-way ANOVA analysis was first conducted and the overall F value and p value examined, then the post hoc pair-wise comparison was run to identify the exact differences. To address RQ5a, the test results were compared across time for each of the two groups. To answer RQ5b, we compared the two groups of students in terms of their performance at the three time points. We expected to find

students in the EG performing similarly to the CG at the beginning but outperforming the CG later because they had received metacognitive instructions.

The interview protocol and researchers' notes were analysed to further explore pertinent influencing factors in response to RQ5c. The interview data were first transcribed into text Microsoft Word files using the auto-translating tool Xunfei translator, and then the researcher double-checked the transcription. The transcriptions were stored online together with the researcher's daily teaching notes using the online documentation tool Shimo Document. Afterwards, these data were analysed using the conceptual coding framework employed in the open-ended question survey data analysis. Details of the findings are given in the next section.

5.4 Findings

This section examines the impact of this intervention over time and compares the similarities and differences between the CG and the EG. First, the two groups' overall improvement across three time points is reported. Second, detailed group improvement over the three time points is also compared and reported. Third, the differences and similarities between the two groups are analysed. Finally, potential influencing factors as observed in the process of intervention are also examined.

5.4.1 Comparison of pre-, post-, and delayed post-test writing performance

Descriptive statistics and one-way ANOVA were computed to compare pre-test results with post- and delayed post-test results to learn about the potential changes in the experimental group and the control group. The effect size (eta square) was also calculated for both groups. Details of the results are shown in Table 5.3 in Appendix I.

Table 5.3 in Appendix I shows the group means and the standard deviations of the overall discourse scores and sub-scores for EG and CG at the three time points; it also shows the significance of the mean values between pre- and post-test and between pre- and delayed post-test within each group. In Table 5.3, the EG showed a significant increase in the mean value of discourse features in the post-test, as represented by [DO], with $F(2, 69) = 9.538$, $p = .000$, and $\eta^2 = .293$. Specifically, there were significant increases on [F1], [F4] and [F6] in the post-test, with $F(2, 69) = 15.519$, $p = .000$, and $\eta^2 = .403$ for [F1], $F(2, 69) = 15.242$, $p = .000$, and $\eta^2 = .389$ for [F4], and $F(2, 69) = 6.787$, $p = .002$, and $\eta^2 = .287$ for [F6]. Surprisingly, however, there was no significant increase in the delayed post-test as compared with the pre-test. Instead, the mean values of discourse features dropped slightly between the post-test and the delayed post-test (2.868 vs 2.385).

The CG also had significant increases in [F1] and [F4] in the post-test, with $F(2, 69) = 6.582$, $p = .004$ and $\eta^2 = .216$ for [F1], $F(2, 69) = 8.906$, $p = .016$ and $\eta^2 = .157$ for [F4]. The CG, however, did not have a significant increase in F6, though it demonstrated a significant increase from the pre-test to the delayed post-test on [DO], with $F(2, 69) = 5.624$, $p = .004$, and $\eta^2 = .200$.

5.4.2 Comparison of writing performance between experimental and control groups

The mean values and one-way ANOVA were computed to detect if there were significant differences between the EG and CG at the three time points. Details are shown in Table 5.3 in Appendix I and Table 5.4 below.

Table 5.4

Comparison of group performance at three time points

Effect	Test time	F	*p*	Eta squared (η^2)
[DO]	Pre-test	.981	.327	.021
	Post-test	9.099	**.004**	.165
	Delayed Post-test	1.278	.264	.027
[F1]	Pre-test	.012	.912	.000
	Post-test	5.689	**.021**	.110
	Delayed Post-test	2.329	.134	.048
[F2]	Pre-test	1.007	.321	.021
	Post-test	.928	.340	.020
	Delayed Post-test	.132	.718	.003
[F3]	Pre-test	.487	.489	.010
	Post-test	4.646	**.036**	.092
	Delayed Post-test	.655	.422	.014
[F4]	Pre-test	.841	.364	.018
	Post-test	11.628	**.001**	.202
	Delayed Post-test	2.589	.114	.053
[F5]	Pre-test	1.267	.266	.027
	Post-test	1.918	**.173**	.040
	Delayed Post-test	.630	.526	.011
[F6]	Pre-test	.345	.560	.007
	Post-test	4.861	**.033**	.096
	Delayed Post-test	.029	.866	.001

(Note: *p* values are two-tailed; [DO]: discourse overall; [F1]: topic/focus; [F2]: thesis statement; [F3]: controlling idea; [F4]: reader orientation; [F5]: body paragraphs; and [F6]: conclusion.)

Table 5.4 shows a comparison of the two groups' performance at the three time points. More visible features of the trends for each subcategory are shown in Graphs 5.1 to 5.7. As shown in Table 5.4, the EG and the CG performed as they had at the beginning, with no significant differences. In the post-test, the EG outperformed the CG significantly in overall discourse

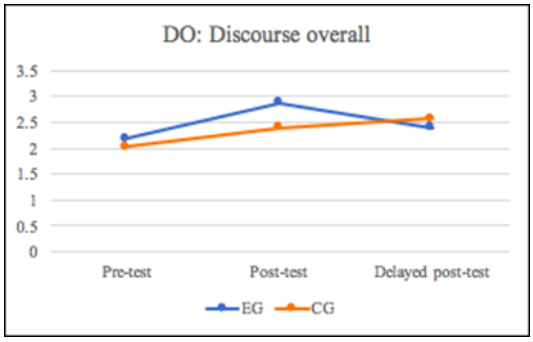
coherence [DO] as well as in the four subcategories, as noted in the bolded cells, with $F(1, 46) = 9.099$, $p = .004$, and $\eta^2 = .165$ for DO, $F(1, 46) = 5.689$, $p = .021$, and $\eta^2 = .110$ for [F1], $F(1, 46)=4.646$, $p = .036$, and $\eta^2 = .092$ for [F3], $F(1, 46) = 11.628$, $p = .001$, and $\eta^2 = .202$ for [F4], and $F(1, 46) = 4.861$, $p = .033$, and $\eta^2 = .096$ for [F6]. No significant difference between the two groups was detected in [F2]. Moreover, the two groups did not have significant differences in their delayed post-test performance.

This evidence suggests that the present intervention was partially significant in improving students' discourse competencies in writing. The present intervention focused on the writing of the introduction (demonstrated through [F4]) and conclusion (demonstrated through [F6]); the specific significance on these two subcategories may also suggest that this intervention had a positive impact. Lack of significant differences between pre-test and delayed post-test in the EG suggests that the impact of the intervention was not stable. Potential influencing factors were reflected towards the end of the findings section.

Also, as shown in Graphs 5.1 through 5.7, the CG showed steady progress across the three time points on all the subcategories except for [F3], although the increases were not statistically significant. In the delayed post-test, the CG had slightly higher mean scores than the EG for almost all the sub-discourse categories, although the difference was not statistically significant. This surprising result warrants further reflection on the intervention process.

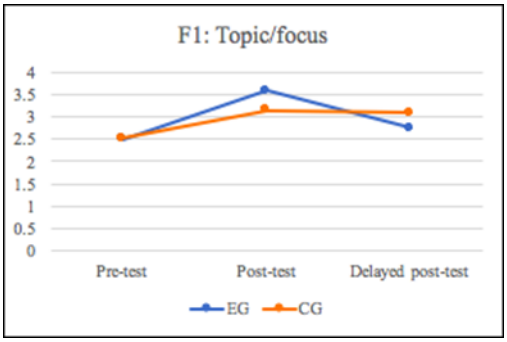
Graph 5.1
DO trend

Graph 5.2
F1 trend



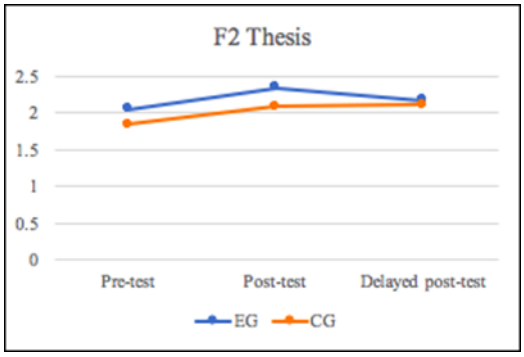
Graph 5.3

F2 trend

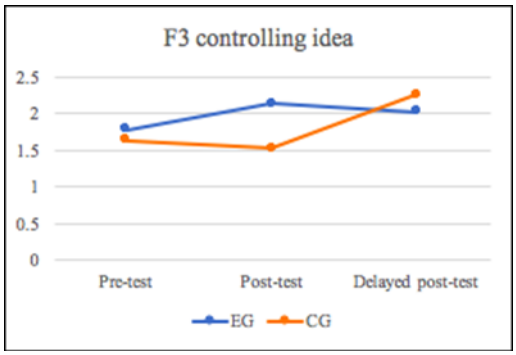


Graph 5.4

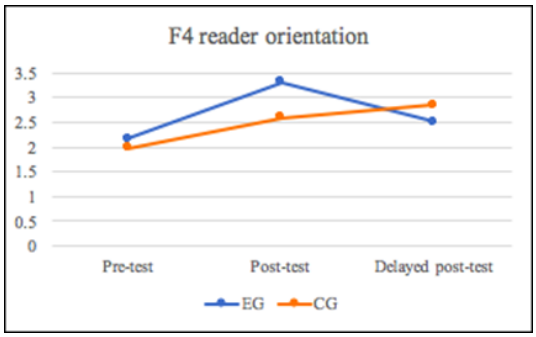
F3 trend

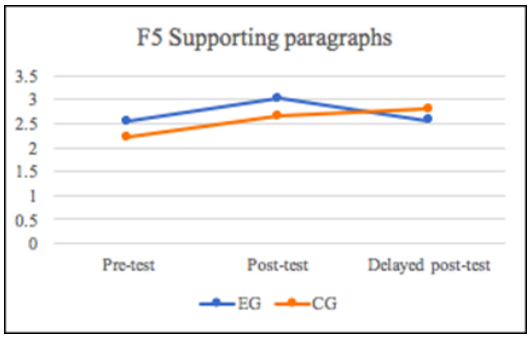


Graph 5.5
F4 trend

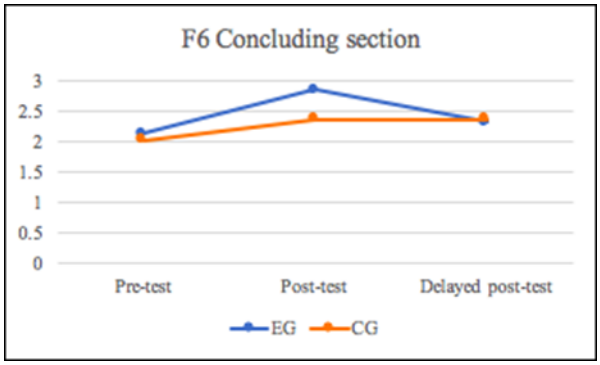


Graph 5.6
F5 trend





Graph 5.7
F6 trend



5.4.3 Potential influencing factors

As shown above, the ANOVA analysis found some significant improvement by the EG in the post-test; however, the CG outperformed the EG in the delayed post-test, though the difference was not statistically significant. This section examines possible influential factors as perceived from the analysis of the qualitative data, including interviews, teaching notes and other related documents collected in the intervention process.

Interview analysis identified two salient benefits students perceived from the intervention activities: an increased familiarity with the strategies and skills of preparing for academic writing and strengthening of students' independent learning knowledge and skills—that is, learner autonomy in writing. For instance, Xi in the EG commented that before the intervention exercise, he did not know where and how to begin with extended academic writing and he would spend considerable time thinking about how to write the introduction; with the intervention, he learnt about the main components of an introduction and how to monitor the whole composing process and orchestrate specific processing strategies. He commented, 'That's really great, and I no longer need to take a long time thinking about how to begin a written work.' Another student, Yu, emphasised her change of learning habits: 'I felt I am a bit different from Year 1 study. In the first year of university study, I did not know how to learn by myself because before university study, all my learning was for exam.

Without exams, I did not know what I should learn. Now I have begun to realise the importance of self-autonomy and independent learning. Over the course, I felt my initiative in learning had gradually strengthened.'

Despite students' acknowledgement of the significance of the instruction on knowledge and skills for processing strategies and global coherence, the intervention did not achieve the expected results. One critical mediating factor may have been the participants' differing attitudes. Observation of students' daily learning activities over the whole semester revealed that there was a remarkable difference between the EG and the CG in participants' attitudes to learning. All through the semester, the 30 students from the CG showed a proactive attitude towards learning and other activities on campus. They were even recommended to be the model class for the grade, from among around 30,000 students in the whole university, to demonstrate class activity at a provincial-level university competition to display the spirit of modern youth. On the other hand, the EC tended to be a class of disengaged students who had little enthusiasm for class assignments or tasks that were not compulsory or gave no course credit. It was later learnt that even at the end of Year 1 study, the monitor (who was responsible for class affairs) had applied to leave this class and was transferring to another major. The major student assistants in the EC also expressed the wish for a re-organisation of classes so that students in the EC might be separated into other classes instead of staying together. Their tutor also showed constant concerns and worries about this class's performance and participation in various activities. Thus, in the intervention process, as in most other activities on campus, many students in the EG very often just did what was required in class by the teacher and showed little interest in applying their learning elsewhere. In contrast, some students from the CG were inquisitive about learning and had taken various opportunities to seek suggestions from teachers or peers.

Another factor contributing to the unstable performance of the EG may be the time points at which the post-test and delayed post-test were conducted. The post-test was conducted around the week when students were carefully preparing for their disciplinary writing task, which accounted for 20% of their final course score, and they needed suggestions for widening their thinking and improving their writing. The delayed post-test, however, took place in the last week of the semester, the task was no longer related to any formal exams and the result would not affect the students' academic scores. Given the general mentality of the EG students, they may not have been fully committed to the delayed post-test writing task.

A third influencing factor may be related to the students' genuine acceptance of the intervention contents. While they claimed that this intervention was meaningful and helpful, they also expressed some difficulty in adjusting to this intervention approach, especially when the intervention was related to their disciplinary learning. When asked about the integration of disciplinary knowledge and writing in the language course, Lei commented that she felt that disciplinary learning and the language course were completely different things. The essay structure had been explained in their disciplinary course, and she would do whatever she was told in the disciplinary course to fulfil the writing requirements. If their subject teacher did not mention some features, she would not care about them because adding these features might not contribute to the final score. Student Yu, who commented on benefits of learning autonomy as quoted previously, said, 'In our regular language course, the focus is all about vocabulary, grammar and sentence patterns. We have difficulty adjusting to this intervention approach. Maybe as time goes by, we will get used to it.'

A fourth influencing factor may be related to the instruction language. Yu again in the EG

commented that they developed a more dynamic understanding of the written text in the intervention process, but the reading materials were difficult for them: 'You see, the first response many students had to the reading materials was to take out their cell phones and take a picture to see the Chinese translation of the reading material. But this teaching method is very meaningful and we may need more time to get used to it.' Besides, the intervention language may affect students' comprehension of the instruction tips. Some students gave feedback that they had difficulty understanding the instruction in English, and so a bilingual instruction method may be preferable.

As shown above, overall, the EG had significantly better performance in the post-test than in the pre-test, but there were no significant differences between the pre-test and delayed post-test. The EG also had significantly better performance than the CG in the post-test, but it was at a similar level in the delayed post-test.

5.5 Discussion

The intervention aimed at developing students' GC, especially in terms of introduction and conclusion writing, through a metacognitive instruction approach. Results of the intervention effect showed that the EG had significantly better performance than the CG in the post-test, but it had an overall decline in the delayed post-test. This section first expands on the unexpected intervention findings with reflection on the overall assessment process and then discusses the usefulness of the metacognitive intervention, specifically the six metacognitive procedures.

In the present intervention, there was a significant impact in the post-test but not in the delayed post-test. Reflection on the post-test and delayed post-test processes revealed that the

writing context around the week of the post-test might have had a great influence on student writing. That is, instruction on strategy and its effects on writing could be context dependent, as suggested in many studies (Johns, 1997; De Silva, 2015). In the score-driven learning context, students' singular motivation in writing was to complete the written task to fulfil the academic credit requirements. Around the week of the post-test, students were also busy preparing for their 500-word disciplinary writing. As the post-test and daily intervention exercises were explicitly directed at helping students handle the 500-word essay writing, students had close simulation of this post-test with their authentic disciplinary writing. However, around the week of the delayed post-test at the end of the semester, students did not have any compulsory writing tasks that were related to their academic credits. In that situation, the attitude to writing emerged as a crucial influencing factor. Findings on pertinent influencing factors showed that there was a salient and constant difference between the EG (disengaged) and the CG (proactive) in their attitude to learning and other activities on campus. The different attitudes were likely to have resulted in different goal-setting strategies, so the unexpected different results between the post-test and the delayed post-test are understandable. Hence, as suggested in Grabe and Kaplan (1996), the context of writing, both situational and social-cultural, should be given sufficient consideration in conducting writing intervention to ensure or facilitate students' setting of adequate goals and maximise the effect in achieving these goals.

Furthermore, student interviews about the intervention process and reflection on the unstable performance between post-test and delayed post-test revealed the significant role of disciplinarity in academic writing instruction in university study. The central role of disciplinarity in undergraduate academic writing has been acknowledged in some studies (Antony, 2018; Fox & Artemeva, 2017). By conducting a set of diagnostic writing

assessments in a Canadian university engineering programme, Fox and Artemeva (2017) proposed a disciplinary, ESP-based approach to diagnosing the need for academic support in disciplinary writing. Regarding the focus of this intervention, global coherence presents the overall conceptual structure of a text appropriate to the target communicative purposes. Meanwhile, the evaluation criteria on the appropriateness of the structure in the target community depends on the existing practices in the target disciplinary community. In the conventional writing programme, writing instructors mostly focused on language issues, scarcely involving skills and knowledge explicitly related to disciplinary conventions or practices. Students in the present intervention also expressed surprise and sometimes doubt about the language teachers' instruction on discipline-related knowledge and skills. Existing studies proposed the concept of *legitimate peripheral participation* (LPP) in learning to write in a situated discourse community (Flowerdew, 2016; Lave & Wenger, 1991). Viewing LPP as a discrete set of actions, students can be instructed to participate in the disciplinary practice to a limited extent and with limited responsibility and eventually develop complicated skills adequate to the disciplinary or professional community. Galloway and Rose (2021) emphasised the importance of academic and language support in the English-medium instruction context and proposed that it may be up to language practitioners to make the links between language and content in meeting students' needs. On the teaching and intervention design part, however, the present research experienced some practical challenges for language/writing teachers to clarify the legitimate peripheral boundaries in disciplinary writing practice. For disciplinarity concerns, in the present research the content teachers expressed that the students have frequent and salient language problems which need language/writing teachers' urgent support; and the content teachers themselves would take responsibility for the content issues. In this situation, the researcher had both time and human resource restrains in designing more content-specific reading materials and writing tasks.

Were the intervention materials more relevant to these students' disciplinary courses and the authentic writing tasks from the content courses, things might have worked out differently. In view of teaching needs and students' doubts concerning the integration of disciplinary knowledge in undergraduate academic writing pedagogy, writing teachers' explicit qualifications to impart proper disciplinary knowledge is necessary to retaining students' confidence in transferring related knowledge and skills across courses. An alternative strategy for effective intervention is better design of meaningful academic writing tasks (such as project-based writing). This may help provide students with better motivation and stimulation in producing a meaningful piece of writing.

A third point is the necessity of an effective strategy to integrate the development of L2 proficiency and the flexible application of coherence principles and processing strategies. As perceived from student interviews in this intervention, EFL learners' weaknesses in grammar and vocabulary may hamper their attempts to comprehend the coherence features in reading materials and to create coherence in their own writing. Lee (2002) made a similar observation. Hence, effective integration of grammar and vocabulary features, instead of exclusive focus on coherence teaching, is needed. Hyon (2016) provides useful insights into concrete procedures for such integration. Specifically, discipline-related vocabulary and sentence patterns may be accumulated in the process of genre analysis. About the present intervention procedures, it is beneficial to integrate learning of a set of discipline-related vocabulary and sentence patterns (especially at the discussing, modelling, and memorising stages).

A fourth concern is the importance of cultivating students' proactive attitudes to writing. In the present intervention, the EG was found to be disengaged in most of their daily learning,

while the CG displayed a saliently constant proactive attitude towards learning, which might be an important factor contributing to their writing performances. In L2 learning, the concept of attitude can be understood in two typical ways: a dimension along specificity and generality and its relevance to L2 achievement (Gardner, 1985). As one important affective dimension of motivational dispositions, attitude has been considered a crucial determinant of successful second language learning (Dörnyei & Ryan, 2015; Gardner, 1985; Yashima et al., 2004) and effective writing (Graham et al., 2007). Graham et al. (2007) suggested three possible paths regarding the relationship between writing attitude and writing achievement. They found that writing attitude can affect writing achievement significantly: students with a positive attitude to writing may invest more effort while those with a negative attitude may invest little energy when required to write. The impact of attitude on writing performance may be greater than that of the actual learning behaviour (Lee, 2013). However, many studies on attitude to writing have focused on early school learners in native English cultures. The present intervention perceived the disengaged attitude from EG mainly through classroom observation and individual interviews. It is significant to conduct larger-scale empirical studies to examine more specific interactions between attitude towards writing (especially learning to write) and writing performance. Furthermore, it is meaningful to develop adequate teaching and learning strategies to cultivate in EFL undergraduates a proactive attitude towards writing.

Also, one significant point of the intervention is that the metacognitive intervention method is helpful in activating students' awareness of independent learning and strengthening their autonomy in learning, as found from student interviews both during and after the intervention. Lee (1998) synthesised five crucial factors for developing learner autonomy: voluntariness, learner choice, flexibility, teacher support, and peer support. First, this

intervention began with its original concern for English writing difficulties in disciplinary writing and developed materials about their disciplinary study. This disciplinary perspective encouraged and sustained students' interest and voluntariness in learning. Second, the major moves in introduction and conclusion writing were introduced to students frequently by comparative analysis of different types of writing. Such a knowledge base provides students with choice in the objectives of writing. By comparative text analysis in groups, teacher and peer support were also provided. Third, students' familiarity with the composing processes and the interaction between typical processing strategies provide the flexibility for them to adapt the writing processes and procedures based on individual needs and interests. In brief, the discipline-situated metacognitive intervention approach can facilitate learner autonomy, especially in undergraduate academic writing practice.

Nevertheless, application of the metacognitive approach to novice L2 learners in academic writing encountered some challenges in balancing instruction on discourse knowledge and guidance on processing strategies. In the extended academic writing context, L2 learners are mostly like new comers in a strange land, as in McCarthy's (1987) term. Both the academic writing process and writing knowledge pose new experience to these novice L2 writers. About the present study, there may be three factors that enable the present metacognitive intervention approach to facilitate learner autonomy in writing: (1) increasing genre awareness for the disciplinary purposes of writing, (2) familiarity with the systemic strategy knowledge expected in academic writing, and (3) scaffolding of constant exercise in relevance to students' disciplinary study.

Altogether, with its systemic integration of academic writing knowledge, processing strategies, and linguistic features, the present metacognitive intervention approach could

facilitate novice EFL writers' development of global coherence in academic writing. Nevertheless, the intervention effect was mediated by learner attitude and motivation to write. In the context of EFL learning in China's universities, to scaffold for students' solid mastery of global coherence and processing strategies, establishing a particular course that focuses on English writing for specific academic purposes—namely, an English Specific Academic Purposes reading and writing course—is of great value.

There are three immediate benefits to such a course. First, a concrete course will ensure aggregation of the sets of academic writing knowledge and strategy systems in the curriculum and syllabus all through the teaching, learning and assessment processes. This enables the intervention objectives to stand out in a more explicit and impressive place. Second, this will provide necessary motivation for students' participation in related writing activities. If taking a course, students would usually achieve certain course credits in the learning process. Third, the establishment of such a course may promote the prestige of writing teachers as officially qualified academic writing professionals in related disciplinary fields and strengthen students' recognition of the instruction contents.

5.6 Conclusion

This intervention study was conducted to explore the impact of metacognitive instruction on the improvement of Chinese EFL undergraduate students' knowledge and skills in writing introductions and conclusions. The evidence showed that the present metacognitive intervention approach could facilitate novice EFL writers' development of global coherence in academic writing. This approach was also beneficial in promoting autonomous learning in writing. However, the effect was limited due to the EG students' attitude towards and inadequate motivation for writing. As a result, this study has proposed the establishment of a

particular course that focuses on writing in English for specific academic purposes to help facilitate EFL undergraduate learners' development of global coherence in English academic writing.

This study had limitations. The first was the limitation of the intervention site. This intervention was conducted with the syllabus of a comprehensive general English course, towards which students conventionally expected less academic and less discipline-related learning activities. A more appropriate site—for instance, an English for specific purpose reading and writing course—may elicit a more significant outcome. The second was the limitation of intervention time. Within the limited semester schedule, there are multiple objectives for the reading and writing course to fulfil. Thus, the time and space left for intervention instruction were very tight; only about seven hours of in-class practice in the present intervention. Future studies may explore the metacognitive intervention method in a discipline-related writing course and through an extended longitudinal study to examine its effect on EFL undergraduates' acquisition of systemic discourse features and reading and writing strategies in writing.

This chapter has reported an intervention study that follows a series of diagnoses as reported by textual analysis in chapter 2, a comparative case study in chapter 3, and a survey study in chapter 4. Based on the findings from the diagnosis and intervention studies, the next chapter (chapter 6) will synthesise the findings from these five studies and further expand on discussion about pedagogical instruction of an EFL or L2 writing course in university study and further research directions.

Chapter 6 Synthesis of Chapter Findings, Discussion and Implications

The previous five chapters reported the findings on students' specific problems and difficulties at the discourse level, from textual features to processing strategies, and to academic writing knowledge status. Based on these findings, seven issues arise as the most salient features: *overall discourse competence*, *global coherence*, *genre knowledge*, *reasoning and argumentation skills*, *research/planning strategies, revision/evaluation* and *lexical complexity*. This chapter synthesizes these findings to draw implications for pedagogical instruction of EFL/L2 writing course in university study.

6.1 Discussion on typical discourse level features

6.1.1 Problems in overall discourse competence

In the present study, Chinse EFL learners' overall development of DC were perceived from four perspectives: quantitative analysis of the written scripts, a comparative case study between Hong Kong and mainland contexts, a questionnaire study, and an intervention study. While similarities across these various perspectives may provide rich evidence in confirming certain features, differences among these approaches shed light on deeper understanding of subtle factors in contributing to these students development of DC in English academic writing.

On average, the EFL learners in the present study demonstrated a basic level of discourse features. Specifically, from quantitative analysis of students written scripts in Chapter 2, the low-level essays and the medium level ones are rather similar on the discourse features. There is also a similar trend in respects of processing strategies and academic writing knowledge status between High and Medium/Low proficiency students from the survey study in Chapter

4. Furthermore, in Chapter 3, the comparative case study showed that regardless of English proficiency levels, the mainland participants demonstrated weaker discourse features and insufficient process strategies than their Hong Kong local peers. From the case study in mainland, even the high proficiency participants demonstrated only a fair level of discourse features. That is, students in the mainland context at all the three levels, high-, medium- and low- proficiency students, were not very good at discourse features; the high-proficiency ones were only relatively better than the low- and medium-proficiency peers. There is still considerable space for students at all levels to improve their discourse level competencies.

6.1.2 Problems in particular discourse categories

6.1.2.1 Problems and difficulties with Global coherence

Global coherence was found to be the most problematic aspect among the five main discourse categories. From quantitative analysis of students written scripts in Chapter 2, *controlling idea* and *conclusion writing* are the weakest aspects and there was significant difference between high proficiency and low proficiency students in *conclusion writing*. In the survey study in Chapter 4, about 8% of high-level students considered idea organization as their weakness, while 15% or more medium- and low-level students considered as their special weaknesses. And genre knowledge mostly remained invisible from students' perception, except for its subcategory of /reference, citation & documentation format/. As discussed in Chapter 5, genre knowledge is closely related to global coherence. This indicates that global coherence may pose greater challenges to low proficiency students, especially in the cases of mainland students' writing.

From the mainland case study in Chapter 3, students often had the feeling of doing 'filling in blanks' exercise in completing related writing assignments. In their disciplinary writing, the

subject teachers tended to focus on the content parts and often provided students with a detailed writing template without giving explanations on structural components. In their writing course offered by language teachers, students were also provided with certain templates and writing feedback were often focused on linguistic aspects. Thus, while many students could produce a seemingly well-structured writing but they were not clear why these elements need to go together, or where an alternative structure makes it better.

6.1.2.2 Problem with local coherence

Local coherence focuses on semantic connection between sentences. From the quantitative analysis of students written scripts in Chapter 2, students demonstrated a medium level of local coherence. The comparative case study in Chapter 3 found the local coherence problems mainly resulted from two factors, namely, students L2 proficiency and unnatural insertion of citations. Their L2 proficiency is insufficient in conveying their ideas precisely and they may frequently have related or coherent ideas, but express the ideas in a clumsy way. The other factor is students' anxiety adding citations to fulfill the requirement for references. Students had difficulty in integrating the information from sources. In relation to the processing process, it may be directly related to their inadequate planning. If they had planned well about the structure, the use of source materials, and the purpose of reading references, they would not have inserted citations mechanically. Understanding the purposes of reading references is a crucial step for these students. In comparison, local coherence was identified as a strength in the findings from Hong Kong case study in Chapter 3, this may be because students in that case enjoyed a higher English proficiency. It may also because the students in Hong Kong case study had much less anxiety about meeting requirements on the numbers of citations used in their writing.

6.1.2.3 Problems with reader-writer interaction

Of the five main discourse categories, reader-writer interaction was also frequently identified as a typical problem for both mainland and Hong Kong participants, particularly in relation to complexity of hedges and boosters. From quantitative textual analysis in Chapter 2, students' use of hedges and boosters were relatively better than global coherence features, but poorer than logical connectives; there were no significant differences between high proficiency and medium/low proficiency students. Students on average demonstrated the balanced use of boosters than hedges, with roughly about one hedge and one booster for every 80 words. But they demonstrated very limited range of specific hedging and booster dictions. Among the prepared set of hedges and boosters, they only employed 32% of the total hedging types and 49% of the total booster types. Their most commonly used hedges were 'may (be)' and 'should' and the most commonly used boosters were modal verbs of 'must' and 'will'. From the survey study in Chapter 4, only a few students (1% of the total) mentioned audience awareness issues, for which 2% of low proficiency students claimed to be their strength and 1% of medium proficiency students claimed to be their weaknesses. In addition, the comparative case study in Hong Kong and mainland showed that students displayed little explicit awareness of using particular linguistic features to convey the reader-writer interaction. That is, all these perspectives of reader-writer interaction features suggested that undergraduate EFL/L2 learners had little audience awareness and much less careful thinking on appropriate use of related words/expressions to effective interaction with target audience. Alternatively, this may suggest that features regarding reader-writer interaction are less significant in L2 undergraduate writing.

6.2 Discussion on typical processing strategies

6.2.1 Typical processing features in the whole composing process

Overall, this study found that the temporal distribution in the composing process may have importance influence on the writing efficiency. From the findings of Hong Kong cases in Chapter 3, the present study suggested that skilled management of available time is more important than the total available time. This confirms with Hayes and Nash (1996) that it is the balance of the composing process rather than attempts to maximize planning that contributes to successful writing. Findings from the survey study in Chapter 4 showed that these students on average were more satisfied with their research/planning activities than the revision activity; and students tended to perceive monitoring the composing process and orchestration of typical processing activities as a weakness. As stated in Roca de Larios et al (2008), the time dimension is essential in the functioning of monitoring process. Thus, these identified processing problems and difficulties may point to the imbalanced time allocation across the composing process. And the unsatisfactory imbalance further underlines the importance of skilled monitoring and collaboration of typical processing activities in the whole composing process. Existing studies (Manchon, 2018; Tillema et al, 2011) suggested that time-on-task condition should be a key variable and be given prominent concern in writing studies in that both strategy use and the linguistic processing in the process of writing are time-dependent. Given these considerations, efficient time management skills among the core composing activities are likely to the first and maybe the foremost component in effective writing.

In addition, the survey study in Chapter 4 identified a high positive correlation between *Research* and *Planning*, and a weak correlation between *Research* and *Revision*/*Evaluation*, with *Planning* having a high correlation with *Formulation*. The core concern on idea-generation may account for the high correlation between *Research*, *Planning* and *Formulation*. This may suggest that *Research* and *Planning* strategies have substantial

influence on the Formulation process especially regarding idea generation/organization; effective research and planning activities can facilitate the formulation process greatly.

6.2.2 Specific processing difficulties in the composing processes

One salient weakness was related to students' insufficiency in academic reading at the research stage. From mainland cases in comparative case study in Chapter 3, in students' daily experience, most of their English reading was about English novels or proses. While reading disciplinary materials in English, their attention was on the content points. The languages they were exposed to were informal and the structure of their daily reading materials were also quite different from academic reading materials. In addition, when reading English academic materials, these EFL learners often relied on machine translation for patch reading and some EFL learners even used machine translation in the formulation stage of their writing to translate their own ideas from mother tongue to English. The trend of EFL/L2 learners' reliance on machine translation for academic reading comprehension and/or formulation of ideas in English writing has been notified in existing studies (Groves & Mundt, 2015; Lin & Morrison, 2021). While the convenience and benefits of machine translation in academic reading makes it difficult to reject this technology, use of such tools is highly problematic especially for students' comprehension of macro-level features such as organization and construction of an academic argument (Groves & Mundt, 2015). Research is essential in academic writing which provides primary sources for problem formulation and idea generation. But as observed in Hamilton (2018), the significance of academic reading in appropriate academic writing is not fully understood. Hence, an early and guided learning of academic reading is necessary.

Goal setting in planning may be another important factor affecting writing strategies and written outcome. From intervention study in Chapter 5, students goal setting in writing vary to the external writing conditions, which may in turn affect the written outcome. As indicated in Cumming et al. (2002), goal setting is valuable in reflecting peoples' learning/writing intentions and helpful for improving peoples' performance as well. On one hand, effective goal setting entails self-reflection and external feedback. On the other hand, adequate goal setting can facilitate more efficient self-monitoring and reception of feedback in moving forwards to achieving these goals. Cumming et al. (2002) also suggested six components to understand the nature of goals, including the goals themselves, the origins of the goals, responsibility for fulfilling the goal, situational conditions, resources for assisting in realizing the goals and strategies for action. It is meaningful to track students' actions on goal setting and goal adjustment throughout the composing process and elicit substantial data for understanding its impact on students' overall writing efficiency.

A third problem is related to deadline fighters. Both in mainland and Hong Kong, our case studies found that students with lower English language proficiency tended to put relatively large portion of time and efforts into research and planning and often had insufficient time for evaluation/revision. The survey study also showed that students perceived relative sufficiency of research and planning but scarcely had mentions of revision. And such tendency subsequently may cause them to fall into rush deadline fighters, reserving little time for evaluation and revision. Inadequate processing knowledge and skills, particularly weak time management skills in monitoring the whole composing process, are the direct factors resulting in students inefficient composing processes.

Another potential factor may be related to students' attitude and emotions. As found in the intervention study, the experimental group and the control group had remarkably different attitudes and emotions to learning at the post-test, which probably accounted for the unstable performance in discourse features by the experiment group. The interaction between attitude and emotions and university student writing was also suggested in Kervin and Barrett (2018). They found that students who used writing assistance heavily often showed procrastination behaviors, with needs on repeated brainstorming and planning instead of actual on-the-task writing. They further suggested that procrastination is related to self-regulation of one's behaviors, thoughts and emotions. Specifically regarding the relevancy of emotions with procrastination behaviors, they proposed mindfulness, i.e., paying focused and non-judgmental attention to the present, to help student writers deal with emotional distress. Kervin and Barrett's (2018) mindfulness treatment strategy can be understood as a specific method of time management skills in monitoring the whole composing process in a particular context. As a whole, instruction on flexible time management skills in monitoring the whole composing process may be a useful approach to handle the deadline fighter phenomenon in undergraduate writing and consequently improve their writing efficiency.

6.3 Discussion on academic writing knowledge status

From the survey study in Chapter 4, the majority of students were more concerned about linguistic skills than other aspects of knowledge components in writing; they were particularly unsatisfied with their disciplinary-related vocabulary size. From the mainland cases in comparative case study in Chapter 3, low proficiency students expressed special concern on vocabulary size to convey their ideas and high proficiency students showed somewhat flexible vocabulary skills to express themselves in different ways. Also, relatively more students were satisfied with their reasoning and argumentation skills but were

unsatisfied with their disciplinary knowledge. There are relatively few mentions of genre knowledge strengths or weaknesses from students' perspective. This result confirms the findings from many other studies that EFL/L2 learners in EMI context were more concerned about word and sentence level features, especially about discipline-specific vocabulary, but showed less concern on discourse level features (Lin & Morrison, 2021).

It has been constantly argued that L2 writers' language proficiency significantly undermines their ability to produce high-quality text (Hinkle, 2011; Qu, 2017). However, it still remains unclear to what extent and under what conditions L2 learners may attain the level of L2 proficiency as expected in producing an effective written product. But many researchers of L2 learning and development have emphasized that even highly educated adult L2 learners may need years of language training to achieve the level of proficiency expected in an effective writing (Hinkel, 2011). Regarding L2 learning in tertiary education, this suggests that the pursuit of high language proficiency before the development of discourse competence may be not practical, especially considering the relatively short span and often highly tight schedule in university study. In other words, the development of L2 proficiency itself seems not a practical solution for promoting the development of discourse competence in L2 writing.

The very limited mentions of genre knowledge on the students' part may suggest that these undergraduate EFL/L2 learners have weak genre awareness. The development of DC is one of the central goals in teaching academic writing in classroom setting. For teaching communicative affairs, Swales (1990) proposed that concentration on genre features may be an effective initiator in helping advance students' communicative competence in that these learners' main interests are often outside the linguistic sciences. The need for writing

curricula to begin with genre, especially for novice students, is also frequently proposed in many other studies (Gardner, 2012; Hyland, 2004; Johns, 2011; Wingate, 2018).

6.4 Discussion on the interactions among discourse level qualities, processing strategies, and academic writing knowledge base

The present study found that the development of DC in undergraduate EFL/L2 learners academic writing is likely to be affected by their processing strategies, L2 proficiency, genre knowledge and disciplinary/domain knowledge. For students' effective development of DC across curriculum, there is a need of collaboration between language teachers, writing instructors and content experts. Particularly, the present study found that efficient collaboration between writing teachers and content experts had significant contribution to smooth the implementation of the intervention process. As admitted in existing study, writing/language teachers' collaboration with content lecturers was often difficult due to the need of long-term partnership and some socialization (Li, 2020; Wingate, 2018). Given the relatively stable teaching staff but tight teaching and learning schedule in the EMI context in a local non-English culture, the present study showed that among the three common types of collaboration, writing teachers' initiative for suggestions on subject-related materials from content lectures works as a practical form to enhance university students' development of academic writing ability (Wingate & Tribble, 2012). Specifically, while designing and developing reading materials for the intervention study, the researcher had in-depth discussion with disciplinary teachers in order to select related sources of references and decision on certain topics; the content teachers provided the first-hand information on the proximity of the appropriateness and difficulty level of potential reading materials in relation to students' disciplinary study.

Meanwhile, the intervention study showed that the metacognitive approach may provide systemic guidance in helping achieve such kind of collaboration. As proposed in Chapter 5, the relationship between discourse level qualities, processing strategies and academic writing knowledge can be perceived as the interrelation as fish, fishing and water. The process-genre method as employed currently in some studies (e.g., Huang & Zhang, 2020) can be conceived as a concrete realization of the metacognitive approach, shedding practical lights on handling the relationship between discourse features and processing strategies in classroom instruction. As conveyed from the ordering of the two key elements in the process-genre method, 'process' strategies may work as the initiating and recursive learning behaviors and the acquisition of 'genre' knowledge as exploration of relevant discourse features in specific social and disciplinary context (Huang & Zhang, 2020; Paltridge, 2012). But Huang and Zhang's (2020) focus was on relatively short argumentative essays for general writing purpose in the college English learning context. That is, the approach proposed in their study did not consider the role of disciplinary context in the intervention procedures. And the genre features in the specific disciplinary community have significant influence on the demonstration of discourse level qualities.

Regarding this concern, the present study adopted a process-genre approach from an ESP perspective to promote the EFL learners' development of DC in university writing. Particularly in the present intervention study, the exercise of certain composing strategies was practiced iteratively in relation to genre analysis of specific discourse features in related texts as selected from discipline-related journal articles, book chapters and magazine essays. The intervention was found effective but mediated by participants' attitude toward writing.

Apart from exploring efforts in maximizing students' abilities in flexible coordination of the processing strategies, linguistic skills and subject knowledge, an additional concern is on students' attitude toward writing. For the participants on the mainland site, their major driving force in writing is score. If no score is provided, most of them would take little effort in learning and writing; the more scores provided, the more efforts they would make. The action of writing itself is the process of thinking and writing can promote rational and scientific thinking (Bean, 2011; Langer & Applebee, 1987; Menary, 2007). But in the present teaching and learning context, without the drive of score or once they have achieved the required score, students would be reluctant to think or be inert in thinking on their own. The overwhelmingly score-driven mind-set may significantly influence their learning and writing strategies, which may have subsequence impact on the written product and on the transferring of related knowledge and skills across courses. As a purposeful selection and organization of facts, opinions, or ideas, writing can be considered as a process of thinking (Arapoff, 1967; Matsuda, 2003). It is suggested that thinking is best taught in relation to content (Langer & Applebee, 1987). Meanwhile, the intervention process in the present study showed that students tended to relate more and show more empathies when there was explicit relevancy between the teaching and learning task and their disciplinary study. Hence, considering both cognitive demands and affective factors, proper integration of discipline-related knowledge is of great significance for classroom instruction in university writing.

6.5 Summary

As a whole, the EFL/L2 learners in university study on average demonstrated basic to fair level of discourse features in English academic writing. One of their crucial problems and difficulties is related to global coherence. The intervention study suggested that a metacognitive approach is effective in facilitating EFL learners' development of global

coherence, albeit mediated by learner attitude. Besides global coherence, these EFL learners also have problems with topic building, complexity of hedges and boosters, and local coherence. As sequencing and priority of learning is a necessary pre-conditional step for classroom instruction, this thesis study suggests that topic building may work as the starting point and priority should be given to global coherence over reader-writer interaction. Regarding the processing strategies, three components in the processing strategies need particular attention: academic reading in research, goal setting in planning, and collaboration in evaluation/revision. In addition, effective coordination and the regulation of the whole composing processes is critical for novice L2 writers, of which flexible time management among different processing strategies may be of great value. In respect of academic writing knowledge base, on the students' part, they perceive discipline-related vocabulary to be their greatest needs; while from the developmental perspective, this study suggests that instruction on genre knowledge in relation to disciplinary background is of great value for classroom academic writing exercise in university study.

6.6 Implications for English for Specific Academic Purpose writing pedagogy and research

The main purpose of the present thesis study was to conduct a systemic diagnostic assessment to identify undergraduate EFL learners' specific problems and difficulties in English academic writing in order to provide meaningful pedagogical instructions to help facilitate these students develop of DC in writing and ultimately improve their writing ability as a whole. In general, evidence from different sources all pointed to the importance of disciplinary contexts in development of EFL learners' DC in academic writing. Hence, the following implications were suggested in relation to teaching, learning and assessment of

discourse related features in the context of an English for Specific Academic Purpose (ESAP) oriented writing course.

6.6.1 Implications for implementation of diagnostic assessment in ESAP writing

The present study suggested the emphasis of diagnostic language assessment in two aspects: the scope of assessment and the manner of assessment. Concerning the content of assessment, the instruments of a diagnostic language assessment can not only aim at diagnosing what has been learnt, but can also have potential in diagnosing specific problems/difficulties in the process of learning for the effect of better improving the learning process. The examination of the composing process in the present study is a meaningful trial in this aspect.

Regarding the manner of assessment, it is valuable to have further exploration on the role of learner-as-assessor in diagnostic assessment. Especially for the complex adventure of extended academic writing, there are diverse contextual factors across time, place and available sources that may affect the writing process and the written product. The validated discourse rubric displayed potential to facilitate this type of diagnosis. The ultimate goal of diagnostic assessment is for the provision of effective treatment methods to help repair or improve the target situation. For diagnostic information to scaffold a meaningful bridge between teaching approaches and learning efficiency, one useful method may be for the researcher to conduct the diagnostic assessment with a collaborative role rather than an evaluative role (O'sullivan, 2012). That is, in the assessment process, the teacher supports students to create a diagnostic environment and students themselves work as the major assessors. In the process of self/peer evaluation, students themselves take the agent role and can have constant practice of the target skills and accumulate related knowledge. In this way, the diagnostic assessment itself becomes a means for promoting learning. Furthermore, if

integrating self/peer evaluation as an operational means for diagnostic assessment, the assessment instrument should also be designed and developed comprehensible to the target test takers to enable their effective application of the instrument.

6.6.2 Implications for pedagogical instructions on development of DC in ESAP writing

The present study initiated a metacognitive method for treating the diagnosed discourse problem. The intervention result showed that a process-genre based approach with an ESP perspective would shed a meaningful and practical set of procedures in developing undergraduate EFL/L2 learners DC in academic writing. Firstly, this approach embraces the core academic writing knowledge component 'genre', which has a dominant influence on effective employment of discourse level features. Equipment with adequate genre knowledge will set a solid foundation for flexible and effective employment of related discourse features in academic writing. Secondly, this approach admits the leading role of processing skills in developing students DC in academic writing. The major goal of university writing is to initiate students into academic discourse community, helping them transfer their learning in language course into their content learning or later professional life. Mastery of proper processing strategies has the potential to scaffold students for acquiring such transferable knowledge and skills. Thirdly, the ESP perspective of writing instruction has the advantage in eliciting students' proactive attitude and emotions and retaining their long-term motivation in the complex and challenging adventure of writing.

6.6.3 Implications for research on ESAP writing strategies

The present study found the EFL learners often had insufficient writing strategies, which may indicate that classroom teachers should give more attention to the writing process. Especially for the classroom instruction of writing strategies on undergraduate EFL learners, the primary

concern is on monitoring of the whole composing processes, and the major attention is better given to the quality of research and planning at the beginning stage. With monitoring strategy, an important perspective is research on monitoring time allocations among typical processing activities. With awareness of the significance of time dimension in monitoring the composing process, students can be encouraged to have more practices on research and planning skills. As students become more skilled and accumulated more knowledge in writing, later emphasis can be put on evaluation and revision strategies. Thus, continuous studies aim at understanding these novice writers' overall composing processes and improving their researching and planning skills may provide meaningful guide for the application of theories into classroom practice.

6.6.4 Implication for research on ESAP writing pedagogies

The present study suggested the necessity and the benefit of establishing an ESAP writing corpus which aims at meeting EFL undergraduates' needs in different disciplinary studies. Further research can be conducted which aims at students' effective use of the corpus for developing their writing ability. Take the present focus group, business undergraduates for instance. This ESP writing corpus can include book, journal articles, and written exercises that are frequently encountered in students' daily disciplinary study. Such corpus would work as useful companion resources both for strengthening students' disciplinary learning and facilitating their development of writing-related abilities. Particularly, students can be guided to refer to certain discourse features under question and make comparative textual analysis between different genres of written scripts.

This chapter synthesized the findings from the four empirical studies, i.e., the textual study, the multiple case study, the survey study, and the intervention study, as reported in previous

four chapters. Further discussions were conducted regarding specific discourse features/problems, typical processing strategies/difficulties, academic writing knowledge status, and the interactions within and among these three dimensions. The next chapter (chapter 7) will summarize the main findings and implications as derived from the present thesis project.

Chapter 7 Conclusion, Limitations and Future Studies

Discourse level qualities pose significant challenges to undergraduate EFL/L2 learners in academic writing. Aiming at precisely analyzing a problem and identifying its causes for the purpose of effective treatment, diagnostic assessment provides meaningful guidance for a clear understanding of these challenges and problems in the expectation of activating effective treatment measures. Embracing diagnosis, treatment, and evaluation as a cycle, this thesis study conducted a systemic diagnostic assessment to help identify EFL/L2 learners' specific problems and difficulties in the development of DC in English academic writing in the EFL/L2 learning context. A three-tiered diagnosis procedure was implemented by means of textual analysis, multiple case study, a comparative study, a survey study and an intervention study. Triangulation of the diagnosis results was explored through analyses of multiple sources of data.

This chapter first reflects on the overall significance of the three-tiered diagnostic assessment procedures. It then addresses the challenges and opportunities as perceived in implementation of this multi-level assessment, i.e., in diagnosing the specific textual problems, in probing into the writing process, and in provision of meaningful treatment. Finally, related limitations are also acknowledged and some plans on further research are contemplated.

7.1 Conclusion

Overall, a substantial set of evidence from the present study showed that the three-tiered diagnostic assessment procedure is effective in identifying specific discourse level problems and difficulties to inform teaching and learning of EFL writing. One core element that contributes to the effectiveness is the shared emphasis on the specific context of learning among the assessment target (DC) and the assessment method (diagnostic assessment). On

one hand, specificity of context is crucial for effective presentation of discourse level features (Bazerman, 1988; Wingate, 2012). On the other hand, for diagnostic assessment to be effective, purposeful and relevant, a defining characteristic is to situate the assessment in specific curriculum aims or pedagogical contexts (Cumming, 2015). Furthermore, the saliency of context is closely related to content domain and/or discipline especially in writing assessment (Grab & Kaplan, 1996; Hamp-Lyons, 1990, 2016). When implementing a diagnostic assessment procedure in university learning, Fox et al (2016) suggested that the consideration of disciplinarity is critical in relating identification of discourse problems to effective intervention. In brief, relating assessment of discourse level qualities to diagnosis for effective treatment, the present study suggested that contextuality, disciplinarity in particular, is a critical element for ensuring effective treatment when initiating a diagnostic assessment especially in relation to academic writing exercise in university study.

In view of the whole diagnosis procedure, the present study suggested that the role of diagnosticians is crucial in diagnosis. As Harding et al. (2014) stated, diagnosticians play a critical role in diagnostic assessment procedures in that they decide when to use the assessment tool and how to use related resources. But based on the literature review of the present study, scarcely any existing studies on diagnostic language assessment had more detailed elaboration on the knowledge and skills as needed on the part of a diagnostician with the only exception of Harding et al. (2014). In their study, Alderson et al. (2015) identified three components that contributed to the knowledge of a diagnostician: good training, availability of resources, and a diagnostician's experience and expertise. In relation to university academic writing, a qualified diagnostician is expected to be equipped with proper academic writing experience, processing knowledge, academic writing knowledge. More importantly, the diagnostician should have clear understanding with the interaction among the

three components so as to make adequate synthesis of various types of evidence for final decision on diagnosis (Harding et al., 2014). In classroom setting, the diagnostician is often the writing teacher. Hence, the on-going teacher training or professional development program may play a significant role in facilitating effective implementation of diagnostic assessment.

Diagnostic assessment also needs the support of proper assessment tools. This thesis project developed a discourse rubric to assess students' specific discourse level problems and difficulties at an early stage of academic writing. In this thesis study, textual analysis of written script, case study results and intervention study showed that the present discourse rubric was learner-friendly and was able to provide students with detailed and concrete feedback in a way they could understand and could act upon. In this sense, there is the potential opportunity for using this rubric as a self-assessment instrument. Meanwhile, there were special challenges in developing a proper rubric for diagnostic assessment purpose, especially in two respects: the construct specificity and the scale description on defining and delineating what is problematic and what is normal in EFL learners' exercise of academic writing in university study. This situation is especially salient regarding the seemingly ubiquitous but elusive concept of 'discourse'. Through theoretical synthesis and empirical validation, the present study improved the specificity of discourse competence in writing by decomposing it into ten features under five main constructs. Regarding scale description, the present study established a series of descriptors at each level by the synthesis of five sources of criteria: theoretical conceptualization in writing assessment, empirical findings from existing research work, expert writing, rater experience and related criteria as commonly accepted in other types of assessments applied to university student writing (e.g., IELTS, and

TOEFL). Evidence for quantitative textual analysis show that the scale descriptors developed this way is capable of differentiating different levels of student writing.

In respects of diagnosing the composing process and underlying knowledge foundation, the present study suggested that the integration of a longitudinal case study can help generate rich information in relation to problematic features and uncover pertinent underlying factors. In the present study, emerging from the case study and intervention process was the score-driven motivation in writing as a salient factor contributing to students' written features and related composing strategies. As a result, in designing and developing intervention materials, the present study tried to sustain students' motivation in writing by integrating discipline-related materials and simulation of students' on-going disciplinary writing assignments as references for practice on textual analysis and writing exercises. The intervention showed that the discipline and on-going writing assessment related activities were in sustaining students' motivation and further eliciting more effective discourse level qualities in writing. Hence, the present study suggested that in the future study of diagnostic assessment, the methods of case study and ethnographic study may be given more attention so as to enrich related diagnosis procedures and profiles.

Furthermore, considering the close interface between diagnosis and intervention, this thesis study implemented a metacognitive invention method to treat the diagnosed discourse problems and difficulties. The outcomes of the intervention supported the diagnostic results, but with one extra factor (attitude) as a significant mediator affecting the intervention effectiveness. As a critical catalyst, attitude is found to have direct influence on overall writing proficiency and shapes writing development (Graham et al., 2007). Thus, in the future design of diagnostic assessment procedures, to increase the efficiency of treatment, it may be

necessary to have a systemic synthesis of related literature and take control of the critical factors that will affect the performance of the assessment domain. Meanwhile, as suggested in Alderson et al (2015), the goal of treatment can be problem management instead of problem solving. That is, the treatment may provide strategies on managing the problems or difficulties, but not immediately remove the problems. In this sense, the metacognitive invention method in the present study is also capable of providing problem managing skills, i.e., adequate composing strategies, to help students cope with their discourse level problems and difficulties.

7.2 Limitations

The present thesis study is limited mainly in five aspects: the homogeneous disciplinary context, the sample size of the participants, the focus on one genre type and on the 500-word discipline-specific essay writing, the relatively short-spanned research time in multiple case studies, and the disciplinary specificity in intervention study.

First is about the participants' disciplinary context. This study mainly focused on the writing features of students in a business program. It may be interesting to examine and compare essays written for a different discipline, such as an engineering or a natural science context, to have a more comprehensive view of Chinese EFL undergraduate writers' development of DC in English academic writing.

Second is about the sample size, especially for case study of business students in Hong Kong. Only a small number of students were found suitable for the present study at the research site in Hong Kong context. Future studies may seek more opportunities to explore undergraduate

business students' development of DC in writing in a wider range of educational institutions in a similar EMI context.

The third limitation is related to the focus of genre type. The present textual analysis only focused on one type of essay writing, and the quantitative analysis of student written scripts was only on one task, i.e., the 500-word discipline-specific essay writing. The somewhat exam-conditioned essay writing context is also more or less different from a 2000-word or so academic essay to be written at home. The genre type and specific writing context may have influenced students' composing strategies and development of certain discourse features in essay writing. Hence, it is necessary to conduct similar quantitative textual analysis on a different topic with the same essay genre under different contexts or with a different genre to achieve a more generalizable view of students' development of DC in English academic writing.

The fourth limitation is related to research time. Due to the limitation of personal energy and time concerns, the present study only explored students' writing problems and difficulties at an early stage of their university study. There are certainly some developmental factors that contributes to individual better performance in essay writing. And it is meaningful to conduct an extended longitudinal study of students' development of DC over a long time such as across the four-year degree study to achieve a deeper understanding of students' progress and problems along the academic writing journey.

The fifth limitation concerns disciplinary specificity in intervention study. The intervention in the present thesis project was conducted within the syllabus of a comprehensive general English course, towards which students conventionally expected less

academic and less discipline-related learning activities. The decision on this syllabus context was mainly restrained by actual availability of such a discipline-specific course in the current research site. A more appropriate site—for instance, an English for specific purpose reading and writing course—would have been better for this intervention study to examine the syllabus-specific influence on DC knowledge and acquisition in academic writing.

7.3 Future studies

7.3.1 Design and development of ESAP writing curriculum

For effective development of DC, the mainland case study and intervention study suggest that there is a necessity on design and development of ESAP curriculum for university writing instruction in the EMI context in a local non-English culture. Compared to the often conventionally practiced EFL writing course for general purpose, the ESAP course may be distinctive in four aspects, including curriculum organization, selection of the learning tasks, methods of assessment, and standards of evaluation (Cumming, 2001). In particular, in such ESAP courses, the core elements of such a curriculum should be prioritized on the development of discourse level qualities, processing strategies and discipline-specific genre studies.

7.3.2 Teacher training on ESAP writing instruction

Teacher training is in great need and also under heavy workload for preparing ESAP writing, especially based on a process-genre approach from an ESP perspective. This suggestion derives from the findings of mainland case study and the analysis of the intervention study. Currently the EFL/L2 writing courses are mostly delivered by language teachers. While the language teachers have rich L2 linguistic knowledges and certain gene knowledge, many of them are not well prepared for skilled delivery of proper processing strategies especially for

academic writing and still hold less familiarity with related disciplinary background knowledge. Teaching of writing is both labor-intensive and time-consuming. To have smooth implementation of the ESAP course in practice, effective teacher training is a pre-condition.

7.3.3 Needs assessment from other stakeholders regarding establishment of ESAP course

This study only focused on writing problems/difficulties on the part of students. To make the ESP writing enterprise move toward a more localized and tailored curriculum, it is necessary to take consideration of needs, wants, interests and goals from all the stakeholders, including teachers, students' parents and families, administrators, the educational goals of the university/institute, or even the nation.

7.3.4 Research on novice L2 writers' attitude to academic writing

The present study found that for the participants on the mainland site, their major driving force in writing is score. If no score is provided, most of them would take little effort in learning and writing; the more scores provided, the more efforts they would make. The overwhelmingly score-driven mind-set may significantly influence their learning/writing strategies and efficacy. It is necessary to conduct in-depth investigation on Chinese EFL/L2 undergraduates' attitude to writing and cultivate in them a positive attitude toward writing itself for the ultimate benefits of substantial intellectual and overall personal development.

7.3.5 Research on diagnostic assessment of L2 writing

Based on the intervention outcomes, the present study suggested further research on the role of contextuality in conducting a genuine diagnostic assessment. As stated by Fox et al (2016), 'a diagnostic assessment procedure cannot be truly diagnostic unless it is linked to feedback, intervention, and support' (p. 45). Linking of assessment to feedback and intervention entails the situation of diagnosis in specific context. Future study by conducting comparative

diagnostic assessment among university students between two or more disciplinary contexts may shed further light on understanding the role of context/discipline in diagnostic assessment of university learning experience.

Altogether, this chapter summarized the main findings, the significance and possible impact and implications of the thesis project. Its limitations were also acknowledged in relation to the research context, sample size, focused genre type and research time. Related future studies were proposed regarding teaching, learning and assessment of English as a foreign language learners' academic writing in university study. The sources of references, some supporting materials and more detailed tables are provided at the end of this thesis, along with the ethics approval document.

References

Alderson, J.C. (2005). *Diagnostic foreign language proficiency: The interface between learning and assessment*. Continuum.

Alderson, J. C., Brunfaut, T., & Harding, L. (2015). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics*, *36*(2), 236-260. https://doi.org/10.1093/applin/amt046

Alderson, J. C., Haapakangas, E. L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2014). *The diagnosis of reading in a second or foreign language*. Routledge.

Anthony, L. (2018). *Introducing English for specific purposes*. New York: Routledge.

Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Arapoff, N. (1967). Writing: A thinking process. *TESOL Quarterly*, *1*(2), 33-39. https://doi.org/10.2307/3585751

Arndt, H. & Janney, R. W. (1983). The duck-rabbit phenomenon: Notes on the disambiguation of ambiguous utterances. In W. Enninger & L. M. Haynes (Eds.), *Studies in language ecology* (pp. 94-115). Franz Steiner Verlag.

Askehave, I., & Swales, J.M. (2001). Genre identification and communicative purpose: A problem and a possible solution. *Applied Linguistics, 22*(2), 195–212. https://doi.org/10.1093/applin/22.2.195

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, *36*(3): 199-223. https://doi.org/10.1080/07370008.2018.1456431

Bachman, L. (1990). *Fundamental considerations in language testing.* Oxford University Press.

Bachman, L., & Palmer, A. (2016). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Foreign Language Teaching and Research Press.

Badger, R., & White, G. (2000). A process genre approach to teaching writing. *ELT Journal*, *54*(2), 153–160. https://doi.org/10.1093/elt/54.2.153

Barton, D. (1994). The social impact of literacy. In L. Verhoeven (Ed.), *Functional literacy: Theoretical issues and educational implications*, *(pp.* 185-197). John Benjamins Publishing Company.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. University of Wisconsin Press.

Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. John Wiley & Sons.

Beauvais, C., Olive, T., & Passerault, J. M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, *103*(2), 415. https://doi.org/10.1037/a0022545

Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, *7*(2), 165-182. https://doi.org/10.1075/ijcl.7.2.02bol

Bowen, N. E. J. A., & Thomas, N. (2020). Manipulating texture and cohesion in academic writing: A keystroke logging study. *Journal of Second Language Writing*, 50, 100773. https://doi.org/10.1016/j.jslw.2020.100773

Brown, H., & Adamson, J. (2012). Localizing EAP in light of the rise of English-medium instruction at Japanese universities. *On Cue*, *6*(3), 5-20. http://jaltcue.org/files/OnCUE/OCJ6.3/OCJ6.3_pp5-20_FA_Brown_Adamson.pdf

Brown, H. D., & Abeywickrama, P. (2013). *Language assessment: Principles and classroom practices*. Pearson Education.

Brown, J. D., & Bailey, K. M. (1984). A categorical instrument for scoring second language writing skills. *Language Learning, 34*(4), 21-38. https://doi.org/10.1111/j.1467-1770.1984.tb00350.x

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230-258. https://doi.org/10.1177/0049124192021002005

Bruce, E., & Hamp-Lyons, L. (2015). Opposing tensions of local and international standards for EAP writing programmes: Who are we assessing for? *Journal of English for Academic Purposes*, *18*, 64–77. https://doi.org/10.1016/j.jeap.2015.03.003

Bruce, I. (2008). *Academic writing and genre: A systematic analysis*. Bloomsbury Publishing.

Bruce, I. (2016). Constructing critical stance in university essays in English literature and sociology. *English for Specific Purposes*, *42*, 13-25. https://doi.org/10.1016/j.esp.2015.10.005

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.

Byrne, D. (2005). Complexity, configurations and cases. *Theory, culture & society, 22*(5), 95-111. https://doi.org/10.1177/0263276405057194

Byrnes, H., & Manchón, R. M. (Eds.). (2014). *Task-based language learning–Insights from and for L2 writing* (Vol. 7). John Benjamins Publishing Company.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1–47. http://www.uefap.com/tefsp/bibliog/canale_swain.pdf

Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E. Soler & M. Jorda (Eds.), *Intercultural language use and language learning* (pp. 41-57). Springer.

Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends. *Language Testing*, *20*(4), 369–383. https://doi.org/10.1191/0265532203lt264oa

Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In F. Bachman, & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing* (pp. 32–70). Cambridge University Press.

Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, *32*(3), 385–405. https://doi.org/10.1177/0265532214565386

Chen, J., & Zhang, L. J. (2019). Assessing student-writers' self-efficacy beliefs about text revision in EFL writing. *Assessing Writing*, *40*, 27-41. https://doi.org/10.1016/j.asw.2019.03.002

Cheng, A. (2016). EAP at the tertiary level in China. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 97–108). Routledge.

Choi, I., & Deane, P. (2021). Evaluating writing process features in an adult EFL writing assessment context: A keystroke logging study. *Language Assessment Quarterly*, *18*(2), 107-132. https://doi.org/10.1080/15434303.2020.1804913

Chong, K. K. Y., Ho, A., Cheung, O., Leung, E., & Clarke, P. (2014). Automated prediction of the written errors of tertiary level ESL and EFL learners. *Journal of Second Language Acquisition and Teaching*, *21*, 84–109. https://journals.librarypublishing.arizona.edu/jslat/article/id/272/

Cilliers, P., & Preiser, R. (2010). *Complexity, difference and identity: An ethical perspective* (Vol. 26). Springer Science & Business Media.

Coffin, C., & Donohue, J. P. (2012). Academic Literacies and systemic functional linguistics: How do they relate?. *Journal of English for Academic Purposes*, *11*(1), 64-75. https://doi.org/10.1016/j.jeap.2011.11.004

Coulthard, M., & Condlin, C. N. (2014). *An introduction to discourse analysis*. Routledge.

Crinon, J., & Legros, D. (2002). The semantic effects of consulting a textual database on rewriting. *Learning and Instruction*, *12*(6), 605-626. https://doi.org/10.1016/S0959-4752(01)00031-7

Cross, J. (2011). Metacognitive instruction for helping less-skilled listeners. *ELT Journal*, *65*(4), 408-416. https://doi.org/10.1093/elt/ccq073

Crossley, S. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, *11*(3), 415-443. https://www.jowr.org/articles/vol11_3/JoWR_2020_vol11_nr3_Crossley.pdf

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An investigation into the multiple ways writers can write successful essays. *Written Communication, 31*(2), 184-214. https://doi.org/10.1177/0741088314526354

Cumming, A. (2001). Learning to write in a second language: Two decades of research. *International Journal of English Studies*, *1*(2), 1-23. https://revistas.um.es/ijes/article/view/48331

Cumming, A. (2015). Design in four diagnostic language assessments. *Language Testing*, *32*(3), 407-416. https://doi.org/10.1177/0265532214559115

Cumming, A., Busch, M., & Zhou, A. (2002). Investigating learners' goals in the context of adult second-language writing. In *New directions for research in L2 writing* (pp. 189-208). Springer.

De Guerrero, M. C., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *The modern language journal, 78*(4), 484-496.

https://doi.org/10.2307/328586

De Silva, R. (2015). Writing strategy instruction: Its impact on writing in a second language for academic purposes. *Language Teaching Research, 19*(3), 301-323. https://doi.org/10.1177/1362168814541738

Ding, A., & Bruce, I. (2017). *The English for academic purposes practitioner*. Palgrave Macmillan.

Doe, C. (2011). The integration of diagnostic assessment into classroom instruction. In D. Tsagari & I. Csepes (Eds.), *Classroom-based language assessment: Language testing and evaluation* (pp.63-76). Peter Lang.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford university press.

Dörnyei, Z. (2017). Conceptualizing learner characteristics in a complex, dynamic world. In L. Ortega & Z.H., Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman (pp.* 79-96). John Benjamins.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes, 6*(1), 3-17. https://doi.org/10.1016/j.jeap.2006.11.005

Evans, S., & Morrison, B. (2010). The first term at university: Implications for EAP. *ELT Journal, 65*(4), 387-397. https://doi.org/10.1093/elt/ccq072

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198-208. https://doi.org/10.1016/j.esp.2011.01.001

Feez, S. (1999). Text-based syllabus design. *Interchange*, (34), 5-11.

Ferris, D. R., & Hedgcock, J. (2014). *Teaching ESL composition: Purpose, process, and practice*. Routledge.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Earlbaum.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365-387. https://doi.org/10.2307/356600

Flowerdew, J. (2016). English for specific academic purposes (ESAP) writing. *Writing & Pedagogy*, *8*(1), 1-4. doi: 10.1558/wap.v8i1.30077

Flowerdew, J. & Peacock, M. (2001). Issues in E. A. P.: A preliminary perspective. In J. Flowerdew, & M. Peacock, (Eds.)*, Research perspectives on English for Academic Purposes* (pp. 8-24). CUP.

Flowerdew, L. (2020). The academic literacies approach to scholarly writing: A view through the lens of the ESP/Genre approach. *Studies in Higher Education, 45*(3), 579-591. https://doi.org/10.1080/03075079.2019.1576165

Fox, J., & Artemeva, N. (2017). From diagnosis toward academic support: Developing a disciplinary, ESP-based writing task and rubric to identify the needs of entering undergraduate engineering students. *ESP Today, 5*(2), 148-171.

Fox, J., Haggerty, J., & Artemeva, N. (2016). Mitigating risk: The impact of a diagnostic assessment procedure on the first-year experience in engineering. In J. Read (Eds.), *Post-admission language assessment of university students* (pp. 43–65). Springer International Publishing.

Galloway, N., & Rose, H. (2021). English medium instruction and the English language practitioner. *ELT Journal*, *75*(1), 33-41. https://doi.org/10.1093/elt/ccaa063

Gao, J. (2007). Teaching writing in Chinese universities: Finding an eclectic approach. *Asian EFL Journal, 20*(2), 285-297.

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning, 24*(5), 471-487. https://doi.org/10.1080/09588221.2011.582687

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.

Gardner, S. (2012). Genres and registers of student report writing: An SFL perspective on texts and practices. *Journal of English for Academic Purposes, 11*(1), 52-63. https://doi.org/10.1016/j.jeap.2011.11.002

Grabe, W., & Kaplan, R. (1996). *Theory and practice of writing: An applied linguistic perspective (Applied linguistics and language study)*. Longman.

Graham, S., Berninger, V., & Fan, W. (2007). The structural relationship between writing attitude and writing achievement in first and third grade students. *Contemporary Educational Psychology, 32*(3), 516-536. https://doi.org/10.1016/j.cedpsych.2007.01.002

Green, C. F., Christopher, E. R., & Mei, J. L. K. (2000). The incidence and effects on coherence of marked themes in interlanguage texts: A corpus-based enquiry. *English for Specific Purposes*, *19*(2), 99–113. https://doi.org/10.1016/S0889-4906(98)00014-3

Green, S. (2013). Novice ESL writers: A longitudinal case-study of the situated academic writing processes of three undergraduates in a TESOL context. *Journal of English for Academic Purposes, 12*(3), 180-191. https://doi.org/10.1016/j.jeap.2013.04.001

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112-121. https://doi.org/10.1016/j.esp.2014.09.001

Guo, L., & Wang, H. (2004). An analysis of the error types and their causes in Chinese students' English writing. *Sino-US English Teaching, 1*(5), 9–18. http://www.davidpublisher.com/index.php/Home/Article/index?id=21682.html

Hacker, D. J., Keener, M. C., & Kircher, J. C. (2009). Writing is applied metacognition. In D. J., Hacker, J., Dunlosky, & A.C., Graesser (Eds.), *Handbook of metacognition in education* (pp. 154-172). Routledge.

Hacker, D. J., Kiuhara, S. A., & Levin, J. R. (2019). A metacognitive intervention for teaching fractions to students with or at-risk for learning disabilities in mathematics. *ZDM, 51*(4), 601-612.

Halliday, M., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hamilton, J. (2018). Academic reading requirements for commencing HE students-Are peer-reviewed journals really the right place to start?. *Student Success*, *9*(2), 73-79. https://unistars.org/papers/STARS2017/01C.pdf

Hamp-Lyons, L. (1990). Second language writing: Assessment issues. In B. Kroll (Eds.), *Second language writing: Research insights for the classroom* (pp. 69–87). Cambridge University Press.

Hamp-Lyons, L. (2016). Purposes of assessment. In D. Tsagari & J. Banerjee (Eds.). *Handbook of second language assessment* (pp. 13-28). De Gruyter Mouton.

Hamp-Lyons, L. (2019). Reflecting on the past, embracing the future. *Assessing Writing*, *42*, 1-8. DOI: 10.1016/j.asw.2019.100423

Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, *32*(3), 317–36. https://doi.org/10.1177/0265532214564505

Harrington, D. (2009). *Confirmatory factor analysis*. Oxford university press.

Harris, Z. (1952). Discourse analysis. *Language*, *28*, 1–30. https://doi.org/10.2307/409987

Hayes, J.P., & Nash, J. G. (1996). On the nature of planning in writing. In M.C. Levi & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 29-56). Routledge.

Hayes, J. R. (2012). Modelling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260

He, S.F. (2003). *Contrastive studies of English and Chinese languages*. Shanghai Foreign Language Education Press.

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning*, *15*(1/2), 29. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.6869&rep=rep1&type=pdf#page=33

Hinkel, E. (2011). *Handbook of research in second language teaching and learning*. Routledge. https://citeseerx.ist.psu.edu/messages/downloadsexceeded.html#page=33

Hirvela, A. (2011). Writing to learn in content areas: Research insights. In R.M., Manchon (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 37-59). John Benjamins Publishing Company.

Hoey, M., & Winter, E. O. (1983). *On the surface of discourse*. Allen & Unwin.

Hong, M., & Xu, Y. (2016). A study on local coherence characteristics of non-English majors' English essays. *TEFLE*, 168, 59–64. https://en.cnki.com.cn/Article_en/CJFDTotal-WYDH201602010.htm

Hu, C., & Li, Y. (2015). Discourse connectives in L1 and L2 argumentative writing. *Higher Education Studies*, *5*(4), 30–41. http://dx.doi.org/10.5539/hes.v5n4p30

Huang, Y., & Zhang, J. L. (2020). Does a process-genre approach help improve students' argumentative writing in English as a foreign language? Findings from an intervention

study. *Reading & Writing Quarterly*, *36*(4), 339–364. https://doi.org/10.1080/10573569.2019.1649223

Hyland, K. (1997). Is EAP necessary? A survey of Hong Kong undergraduates. *Asian Journal of English Language Teaching*, *7*(2), 77-99. https://www.cuhk.edu.hk/ajelt/vol7/art5.htm

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing, 12*(1):17–29. https://doi.org/10.1016/S1060-3743(02)00124-8

Hyland, K. (2004). *Genre and second language writing*. University of Michigan Press.

Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2005b). Representing readers in writing: Student and expert practices. *Linguistics and Education, 16*(4), 363-377. https://doi.org/10.1016/j.linged.2006.05.002

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148-164. https://doi.org/10.1016/j.jslw.2007.07.005

Hyland, K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*, *1*(1), 5–22. https://doi.org/10.1075/etc.1.1.03hyl

Hyland, K. (2011). Academic discourse. In K. Hyland, & B. Paltridge. (Eds.), *Continuum companion to discourse analysis* (pp. 171-184). Continuum.

Hyland, K. (2012). Undergraduate understandings: Stance and voice in final year reports. In K. Hyland & C.S. Guinda (Eds). *Stance and voice in written academic genres* (pp. 134–150). Palgrave Macmillan.

Hyland, K. (2013). Faculty feedback: Perceptions and practices in L2 disciplinary writing. *Journal of Second Language Writing*, *22*(3), 240-253.

https://doi.org/10.1016/j.jslw.2013.03.003

Hyland, K. (2015). Teaching writing: Understanding texts, writers and readers. In TESOL Arabia International Conference. Dubai, United Arab Emirates.

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, *6*(2), 183–205. https://doi.org/10.1016/S1060-3743(97)90033-3

Hymes, D. (1967). Models of the interaction of language and social setting. *Journal of Social Issues*, *23*(2), 8-28. https://doi.org/10.1111/j.1540-4560.1967.tb00572.x

Hymes, D. (1972). On communicative competence. In J.B. Pride, & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, *30*(4), 693-722. https://doi.org/10.2307/3587930

Hyon, S. (2016). Using genre analysis to teach writing in the disciplines. In Discipline-specific writing (pp. 91-108). Routledge.Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B.. *Testing ESL composition: A practical approach*. Newbury House.

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B.. (1981). *Testing ESL composition: A practical approach*. Newbury House.

Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for Fusion Model application to LanguEdge assessment. *Language Testing*, *26*(1), 31–73. https://doi.org/10.1177/0265532208097336

Johns, A. M. (1997). *Text, role and context: Developing academic literacies*. Cambridge University Press.

Johns, A. (2008). Genre awareness for the novice academic student: An ongoing quest. *Language Teaching, 41*(2), 237-252.

https://doi.org/10.1017/S0261444807004892

Johns, A. M. (2011). The future of genre in L2 writing: Fundamental, but contested, instructional decisions. *Journal of Second Language Writing, 20*(1), 56-68. https://doi.org/10.1016/j.jslw.2010.12.003

Juarrero, A. (2000). Dynamics in action: Intentional behavior as a complex system. *Emergence*, 2(2), 24-57. https://doi.org/10.1207/S15327000EM0202_03

Jumaat, N.F. & Tasir, Z. (2016). A framework of metacognitive scaffolding in learning authoring system through Facebook. Journal of *Educational Computing. Research, 54*(5), 619–659. https://doi.org/10.1177/0735633115627824

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*(1).

Kervin, C. E., & Barrett, H. E. (2018). Emotional management over time management: Using mindfulness to address student procrastination. *WLN: A Journal of Writing Center Scholarship, 42*(9-10), 10-18.

Kim, Y. H. (2011). Diagnosing EAP writing ability using the reduced reparameterized unified model. *Language Testing*, *28*(4), 509–541. https://doi.org/10.1177/0265532211400860

Kim, Y. H. (2019). An investigation into the dimensional structure of ESL academic writing skills on TOEFL iBT independent essays. *Journal of Asia TEFL*, *16*(1), 307. https://doi.org/10.1177/0265532211400860

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Knoch, U. (2009). *Diagnostic writing assessment: The development and validation of a rating scale*. Peter Lang.

Knoch, U. (2011). Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from? *Assessing Writing*, *16*(2), 81–96. https://doi.org/10.1016/j.asw.2011.02.003

Knoch, U., Elder, C., & O'Hagan, S. (2016). Examining the validity of a post-entry screening tool embedded in a specific policy context. In J. Read (Ed.). *Post-admission language assessment of university students* (pp. 23–42). Springer.

Langer, J. A., & Applebee, A. N. (1987). *How writing shapes thinking: A study of teaching and learning*. NCTE Research Report.

Larsen-Freeman, D. (2017). Complexity theory: The lessons continue. In L. Ortega & Z.H., Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 11-50). John Benjamins.

Lasagabaster, Doiz, Lasagabaster, David, & Doiz, Aintzane. (2021). *Language use in English-medium instruction at university: International perspectives on teacher practice*. Routledge.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Lee, I. (1998). Supporting greater autonomy in language learning. *ELT Journal, 52*(4), 282-290. https://doi.org/10.1093/elt/52.4.282

Lee, I. (2002). Teaching coherence to ESL students: A classroom inquiry. *Journal of Second Language Writing*, *11*(2), 135–159. https://doi.org/10.1016/S1060-3743(02)00065-6

Lee, J. (2013). Can writing attitudes and learning behavior overcome gender difference in writing? Evidence from NAEP. *Written Communication, 30*(2), 164-193. https://doi.org/10.1177/0741088313480313

Lee, J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays.

Journal of Second Language Writing, *33*, 21–34. https://doi.org/10.1016/j.jslw.2016.06.004

Lee, S. (2020). 'The impact of using machine translation on EFL students' writing'. *Computer Assisted Language Learning, 33*(2), 1–19. https://doi.org/10.1080/09588221.2018.1553186

Lee, Y. J. (2021). Still taboo? Using machine translation for low-level EFL writers. *ELT Journal, 75*(2), 1-10. https://doi.org/10.1093/elt/ccab018

Leech, G. (1983). *Principles of pragmatics*. Routledge.

Leki, I. (2007). *Undergraduates in a second language: Challenges and complexities of academic literacy development*. Routledge.

Li, D. C. (2017). *Multilingual Hong Kong: Languages, literacies and identities.* Springer International Publishing.

Li, Y. (2020). Language–content partnership in higher education: Development and opportunities. *Higher Education Research & Development*, *39*(3), 500-514. https://doi.org/10.1080/07294360.2019.1685947

Liang, M. C. (2006). A study of coherence in EFL learners' written production. *Modern Foreign Languages*, 284–292. https://en.cnki.com.cn/Article_en/CJFDTotal-XDWY200603008.htm

Lillis, T. &Scott, M. (2007). Defining academic literacies research: issues of epistemology, ideology and strategy. *Journal of Applied Linguistics, 4*(1), 5-32. https://doi.org/10.1558/japl.v4i1.5

Lin, H.F., & Morrison, B. (2021). Challenges in academic writing: Perspectives of Engineering faculty and L2 postgraduate research students. *English for Specific Purpose, 63*: 59-70. https://doi.org/10.1016/j.esp.2021.03.004

Linacre, J. M. (2013). *A user's guide to FACETS Rasch-Model computer programs program manual 3.70. 0*. Winsteps.com.

Llosa, L., Beck, S. W., & Zhao C G. (2011). An investigation of academic writing in secondary schools to inform the development of diagnostic classroom assessments. *Assessing Writing, 16*, 256－273. https://doi.org/10.1016/j.asw.2011.07.001

Lockwood, J. (2013). The Diagnostic English Language Tracking Assessment (DELTA) writing project: A case for post-entry assessment policies and practices in Hong Kong universities. *Papers in Language Testing and Assessment*, *2*(1), 30–49. http://www.altaanz.org/uploads/5/9/0/8/5908292/3_lockwood.pdf

Lockwood, J. (2016). Towards a specific writing language assessment at Hong Kong universities. In J. Flowerdew, & T. Costley (Eds.), *Discipline-specific writing: Theory into practice* (pp. 196–216). Routledge.

Lv, S.X. (1977). *Study grammar through comparison* (Tong Guo Dui Bi Yan Jiu Yu Fa). Language Teaching and Research Press.

MacArthur, C. A. (2018). Evaluation and revision. In S., Graham & Hebert, M. (Eds.), *Best practices in writing instruction* (pp. 287-308). The Guilford Press.

MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction?. *Reading and Writing, 32*(6), 1553-1574. https://link.springer.com/article/10.1007/s11145-018-9853-6

Macbeth, K. P. (2010). Deliberate false provisions: The use and usefulness of models in learning academic writing. *Journal of Second Language Writing, 19*, 33–48. https://doi.org/10.1016/j.jslw.2009.08.002

Manchón, R. M. (2011). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 61–82). Benjamins.

Manchón, R. M. (2018). Past and future research agendas on writing strategies: Conceptualizations, inquiry methods, and research findings. *Studies in Second Language Learning and Teaching, 8*(2), 247-267.

Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 15-34). Cambridge University Press.

Mauranen, A. (1996). Discourse competence-evidence from thematic development in native and non-native texts. In E. Ventola & A. Mauranen (Eds), *Pragmatics and Beyond New Series*, 195–230.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. PMID: 23092060

McNamara, D. S., Graesser, A., Cai, Z., & Dai, J. (2013, January 1). Coh-Metrix common core T.E.R.A. version 1.0. Retrieved October 26, 2020, from http//:coh-metrix.commoncoretera.com.

McNamara, T., Morton, J., Storch, N., & Thompson, C. (2018). Students' accounts of their first-year undergraduate academic writing experience: Implications for the use of the CEFR. *Language Assessment Quarterly, 15*(1), 16-28. https://doi.org/10.1080/15434303.2017.1405420

Menary, R. (2007). Writing as thinking. *Language Sciences, 29*(5), 621-632. https://doi.org/10.1016/j.langsci.2007.01.005

Miller, R. T., & Pessoa, S. (2016). Where's your thesis statement and what happened to your topic sentences? Identifying organizational challenges in undergraduate student

argumentative writing. *TESOL Journal*, *7*(4), 847–873. https://doi.org/10.1002/tesj.248

Mohan, B., Leung, C., & Slater, T. (2010). Assessing language and content: A functional perspective. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 217-240). Multilingual Matters.

Nesi, H, & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.

Ong, J. (2011). Investigating the use of cohesive devices by Chinese EFL learners. *The Asian EFL Journal Quarterly*, *11*(3), 42–65. https://www.asian-efl-journal.com/main-editions-new/investigating-the-use-of-cohesive-devices-by-chinese-efl-learners/

O'sullivan, B. (2012). Assessment issues in languages for specific purposes. The *Modern Language Journal, 96,* 71-88. https://doi.org/10.1111/j.1540-4781.2012.01298.x

Ouyang, H., & Tang, S. (2006). Authorial identity of Chinese college English learners in argumentative writings. *Journal of PLA University of Foreign Languages*, *29*(2), 49-53.

Paltridge, B. (2012). *Discourse analysis: An introduction*. Bloomsbury Publishing.

Peloghitis, J., & Ferreira, D. (2018). Examining the role of model-texts in writing instruction. *Accents Asia, 10*(1), 17-26.

Pérez-Llantada, C., & Swales, J. M. (2016). English for academic purposes. In E. Hinkle (Eds.), *The handbook of research in second language teaching and learning* (pp. 42-55). Routledge.

Pessoa, S., Miller, R. T., & Kaufer, D. (2014). Students' challenges and development in the transition to academic writing at an English-medium university in Qatar. *International*

Review of Applied Linguistics in Language Teaching, 52(2), 127-156. https://doi.org/10.1515/iral-2014-0006

Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning, 61*(4): 993–1038. https://doi.org/10.1111/j.1467-9922.2011.00663.x

Plonsky, L. (2019). Language learning strategy instructing: Recent research and future directions. In A.U. Chamot & V. Harris (Eds), *Learning strategy instruction in the language classroom: Issues and implementation* (pp.3-21). Multilingual Matters.

Prior, P. (2004). Tracing process: How texts come into being. In C. Bazerman & P. Prior (Eds.), *What writing does and how it does it: An introduction to analyzing texts and textual practices* (pp.167-200). Lawrence Erlbaum Associates, Inc.

Purpura, J. E. (2008). Assessing communicative language ability: Models and their components. *Encyclopedia of Language and Education*, *7*, 53–68.

Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, *38*(3), 444–456. https://doi.org/10.1016/j.system.2010.06.012

Qin, L., & Zhang, J. L. (2019). English as a foreign language writers' metacognitive strategy knowledge of writing and their writing performance in multimedia environment. *Journal of Writing Research*, *11*(2), 394-413.

Qu, W. (2017). For L2 writers, it is always the problem of the language. *Journal of Second Language Writing, 38*, 92-93. https://doi.org/10.1016/j.jslw.2017.10.007

Racelis, J. V., & Matsuda, P. K. (2013). Integrating process and genre into the second language writing classroom: Research into practice. *Language Teaching*, *46*(3), 382–393. https://doi.org/10.1017/S0261444813000116

Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, *7*(3), 180–190. https://doi.org/10.1016/j.jeap.2008.02.001

Read, J. (2016). Some key issues in post-admission language assessment. In *Post-admission language assessment of university students* (pp. 3–20). Springer.

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing, 17*(1), 30-47. https://doi.org/10.1016/j.jslw.2007.08.005

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. Equinox.

Rose, H., Curle, S., Aizawa, L., & Thompson, G. (2019). 'What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation'. *Studies in Higher Education, 45*(11), 2149-2169. https://doi.org/10.1080/03075079.2019.1590690

Ruan, Z. (2014). Metacognitive awareness of EFL student writers in a Chinese ELT context. *Language Awareness, 23*(1-2), 76-91. https://doi.org/10.1080/09658416.2013.863901

Ruan, Z. & Chen, J. (2017). Students' perceptions and practices in L2 disciplinary writing at an English medium university in Mainland China. In H., Reinders, D., Nunan & B., Zou (Eds), *Innovation in language learning and teaching: The case of China* (pp. 197-222). Palgrave Macmillan.

Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. The Guilford Press.

Sang, Y. (2017). Investigate the' Issues' in Chinese Students' English Writing and Their' Reasons': Revisiting the Recent Evidence in Chinese Academia. *International Journal of Higher Education*, *6*(3), 1-11. https://eric.ed.gov/?id=EJ1142300

Sato, M., & Loewen, S. (2018). Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning, 68*(2), 507-545. https://doi.org/10.1111/lang.12283

Shi, L., Baker, A., & Chen, H. (2019). Chinese EFL teachers' cognition about the effectiveness of genre pedagogy: A case study. *RELC Journal, 50*(2), 314-332. https://doi.org/10.1177/0033688217716506

Silva, T., Brice, C., Kapper, J., Matsuda, P. K., & Reichelt, M. (2001). Twenty-five years of scholarship on second language composing processes: 1976-2000. *International Journal of English Studies*, *1*(2), 211-240. https://revistas.um.es/ijes/article/view/48311

Saville‐Troike, M. (1984). What really matters in second language learning for academic achievement?. *TESOL quarterly*, *18*(2), 199-219. https://doi.org/10.2307/3586690

Shepard, C., & Morrison, B. (2021). Challenges of English-medium higher education: The first-year experience in Hong Kong revisited a decade later. In D. Lasagabaster & A., Doiz (Eds.), *Language Use in English-medium Instruction at University* (pp. 167-192). Routledge.

Snow, C. & Uccelli, P. (2009). The challenge of academic language. In D. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge University Press.

Stapleton, P. (2002). Critical thinking in Japanese L2 writing: Rethinking tired constructs. *ELT Journal*, *56*(3), 250-257. https://doi.org/10.1093/elt/56.3.250

Stapleton, P. (2010). Writing in an electronic age: A case study of L2 composing processes. *Journal of English for Academic Purposes, 9*(4), 295-307. https://doi.org/10.1016/j.jeap.2010.10.002

Street, B. V., & Street, B. B. (1984). *Literacy in theory and practice* (Vol. 9). Cambridge University Press.

Susser, B. (1994). Process approaches in ESL/EFL writing instruction. *Journal of Second Language Writing, 3*(1), 31-47. https://doi.org/10.1016/1060-3743(94)90004-3

Swales, J.. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. M. (2012). A text and its commentaries: Toward a reception history of' Genre in three traditions'(Hyon, 1996). *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, (24), 103-115.

Tardy, C. M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of Second Language Writing*, *15*(2), 79–101. https://doi.org/10.1016/j.jslw.2006.04.003

Tardy, C. M. (2009). *Building genre knowledge*. West Lafayette, IN: Parlor Press.

Tardy, C. M., Sommer-Farias, B., & Gevers, J. (2020). Teaching and researching genre knowledge: Toward an enhanced theoretical framework. *Written Communication*, *37*(3), 287-321. https://doi.org/10.1177/0741088320916554

Tardy, J. M. & Swales, J. M. (2009). Form, text organization, genre, coherence, and cohesion. In C. Bazerman (Ed.), *The handbook of research on writing: History, society, school, individual, text* (pp. 565–582). Lawrence Erlbaum Associates.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research, 1*(1), 77-100. https://doi.org/10.1177/1558689806292430

Teng, L. S., & Zhang, L. J. (2016). A questionnaire‐based validation of multidimensional models of self‐regulated learning strategies. *The Modern Language Journal, 100*(3), 674-701. https://doi.org/10.1111/modl.12339

Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference?. *Journal of Second Language Writing, 48*, 100701. https://doi.org/10.1016/j.jslw.2019.100701

Thompson, I. (1986). Readability beyond the sentence: Global coherence and ease of comprehension. *Journal of Technical Writing and Communication*, *16*(1), 131–140. https://doi.org/10.2190/6J1F-DATG-1275-JTFK

Tillema, M., Van Den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2011). Relating self-reports of writing behaviour and online task execution using a temporal model. *Metacognition and Learning*, *6*(3), 229-253. DOI 10.1007/s11409-011-9072-x

Toraskar, H. B., & Lee, P. K. L. (2016). Hong Kong undergraduate students' academic writing: 21st century problems, solutions and strategies. *The Journal of Asia TEFL*, *13*(4), 372-380.

Urry, J. (2005). The complexity turn. *Theory, culture & society, 22*(5), 1-14. https://doi.org/10.1177/0263276405057188

Van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective* (Vol. 3). Springer Science & Business Media.

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Wenden, A. L. (1998). Metacognitive knowledge and language learning. *Applied linguistics, 19*(4), 515-537. https://doi.org/10.1093/applin/19.4.515

Wengelin, Å., Frid, J., Johansson, R., & Johansson, V. (2019). Combining keystroke logging with other methods: Towards an experimental environment for writing process research. In E. Lindgren & Kirk P.H. Sullivan (Eds), *Observing Writing: Insights from keystroke logging and handwriting* (pp. 30-49). Brill.

Widdowson, H. G. (2001). Communicative language testing: The art of the possible. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 12–21). Cambridge University Press.

Widdowson, H. G. (2007). *Discourse analysis* (Vol. 133). Oxford University Press.

Wingate, U. (2012). Using academic literacies and genre-based models for academic writing instruction: A 'literacy' journey. *Journal of English for Academic Purposes, 11*(1), 26-37. https://doi.org/10.1016/j.jeap.2011.11.006

Wingate, U. (2018). Academic literacy across the curriculum: Towards a collaborative instructional approach. *Language Teaching*, *51*(3), 349-364. https://doi.org/10.1017/S0261444816000264

Wingate, U., & Tribble, C. (2012). The best of both worlds? Towards an English for Academic Purposes/Academic Literacies writing pedagogy. *Studies in Higher Education, 37*(4), 481-495. https://doi.org/10.1080/03075079.2010.525630

Xie, Q. (2017). Diagnosing university students' academic writing: Is Cognitive Diagnostic Modeling the way forward? *Educational Psychology*, *37*(1), 26–47. https://doi.org/10.1080/01443410.2016.1202900

Xie, Q. (2020). Diagnosing linguistic problems in English academic writing of university students: An item bank approach. *Language Assessment Quarterly*, *17*(2), 183–203. https://doi.org/10.1080/15434303.2019.1691214

Xie, Q., & Lei, Y. (2021). Diagnostic assessment of L2 academic writing product, process and self-regulatory strategy use with a comparative dimension. *Language Assessment Quarterly*, 1–33. https://doi.org/10.1080/15434303.2021.1903470

Xu, F. (2015). Review on academic writing research. *Foreign Language Teaching and Research, 47*(1), 97-105.

Yashima, T., Zenuk‐Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning, 54*(1), 119-152. https://doi.org/10.1111/j.1467-9922.2004.00250.x

Yang, L. (2006). Nine Chinese students writing in Canadian university courses. In A. H. Cumming (Eds.), *Goals for academic writing: ESL students and their instructors* (pp. 73-89). John Benjamins Publishing.

Ye, W. (2017). Overview of academic writing in China. *NYS TESOL Journal*, *4*(1), 66-73. http://journal.nystesol.org/january2018/8WeierYe.pdf

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. https://doi.org/10.1016/j.jslw.2003.11.001

Yu, A. (2012). Analysis of the problems of the Chinese college students' EFL classroom writings. *International Education Studies, 5*(5), 199–203. https://eric.ed.gov/?id=EJ1067069

Yu, S., Jiang, L., & Zhou, N. (2020). The impact of L2 writing instructional approaches on student writing motivation and engagement. *Language Teaching Research*, 1362168820957024.

Yule, G. (2006). *The study of language* (3rd. ed.). Cambridge University Press.

Zhang, J., & Chen, Y.Q. (1981). *A comparative analysis of grammar between English and Chinese*. The Commercial Press.

Zhang, L. J. (2010). A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Quarterly, 44*, 320–353. https://doi.org/10.5054/tq.2010.223352

Zhang, W.W., Zhang, D.L., & Zhang, L.J. (2021). Metacognitive instruction for sustainable learning: learners' perceptions of task difficulty and use of metacognitive strategies in completing integrated speaking tasks. *Sustainability, 13*, 6275. https://doi.org/10.3390/su13116275

Zhao, C. G. (2019a). Writer background and voice construction in L2 writing. *Journal of English for Academic Purposes*, *37*, 117-126. https://doi.org/10.1016/j.jeap.2018.11.004

Zhao, C. G. (2019b). Chinese EFL learners' perception of the construct of English academic writing. In R. Damerow & K. Bailey (Eds.), *Chinese-speaking learners of English: Research, theory, and practice* (pp.81-95). Routledge.

Zhao, C. G., & Liao, L. (2021). Metacognitive strategy use in L2 writing assessment. *System, 98*, 102472. https://doi.org/10.1016/j.system.2021.102472

Zhao, C. G. & Lyu, Y. (2019). The construct of EFL academic writing: From the perspective of Chinese undergraduates of English. *Foreign Languages and Their Teaching, 309* (6), 69-79. DOI: 10.13458/j.cnki.flatt.004635

Appendix A: A brief overview of the discourse rubric

	Scale		5 Excellent	4 Good	3 Fair	2 Basic	1 Poor
Topic building	Topic/focus	[F1]	A thesis is established in concisely and consistently. The language reads clear.	A thesis is established in concise and consistent terms, indicating the specific focus of the essay.	A thesis is established, but it seems not consistent or somewhat general.	A thesis is established or identifiable, but is vague or unclear. The thesis is not consistent, somewhat shifting.	It is difficult to identify a thesis, focus or the key topic.
	Thesis statement	[F2]	The thesis statement appears at a proper place and is clearly presented. The language reads clear.	The thesis statement is specific. It appears at a proper place and is clearly presented.	The thesis statement appears, but not at a proper place, or it is not clearly stated.	The thesis statement appears somewhat abrupt, or at an improper location. Or a thesis statement seems present, but is not clearly/precisely stated.	If [F1] is '1', [F2] is also '1'; or there is lack of clear thesis statement.
Global coherence	Controlling idea	[F3]	The controlling sentence is concisely presented. The language reads clear.	The controlling idea shows explicitly the overall communicative intent. It is concisely presented.	There is a controlling idea, but it is not in a proper place, or it is not very clearly stated.	There is some sign of the controlling idea, but this idea is vague. Or the controlling idea appears somewhat abrupt and is not very clearly stated.	There is lack of a controlling idea.
	Reader orientation	[F4]	The opening paragraph(s) provide(s) proper reader orientation. The language is fluent and written with proper complexity.	The opening paragraph(s) provide(s) proper reader orientation. The language needs further improvement.	There is some reader orientation. The idea can be sensed but the language somewhat reads difficult.	There is some reader orientation, only by briefly suggesting the context or by directly announcing the topic. Or it includes unnecessary/redundant details.	The opening paragraph(s) provide(s) little or no reader orientation.
	Body paragraphs	[F5]	The body paragraphs have clear and logically sequenced topic sentences. The language is fluent and written with proper complexity.	The body paragraphs begin with adequate topic sentences and provide a coherent series of arguments.	The body paragraphs begin with certain topic sentence. But some topic sentences seem not coherent.	The body paragraphs do not always include a topic sentence. There seems like a topic sentence, but the supporting details did not convey the topic fully.	The topic sentences are often missing or digress from the theme.

	Conclusion	[F6]	The conclusion rounds off the essay effectively. The language is fluent and written with proper complexity.	The conclusion provides a clear (re)statement of the writer's position, summarizes the main points and rounds off the essay effectively.	The conclusion provides limited summary of the main points. Or it lacks some explicit moves. The language somewhat reads difficult.	The conclusion is present but provides limited summary of main points. Or there is some new/redundant information, somewhat unrelated to the conclusion.	The conclusion fails to provide a (re)statement of the main points and does not create a sense of closure.
Local coherence	Theme-Rheme development	[F7]	The Theme-Rheme development is competent The language is fluent and written with proper complexity.	There is competent Theme-Rheme development among sentences.	The local coherence within a paragraph is often perceived in mother tongue. Or coherence break/unrelated idea progression is detected.	There are some unrelated thematic progressions or coherence breaks within a paragraph (which may suggest misplaced information, irrelevant information, or idea in need of more explicit transitional expressions).	There are frequent unrelated thematic progressions or coherence breaks
Logical connectives	Connective complexity	[F8]	The connectives are used varied and carefully, not mechanically.	There is a (wide) range of connectives, with some category including 3 or more types of words.	There is frequent use of some simple connectives, and may be mechanical.	There is certain range of connectives.	There is very limited range or few uses of connectives.
	Connective accuracy	[F9]	The connectives are used precisely and concisely.	The use of connectives fits the grammatical and semantic context. But occasional overuse of certain simple connectives	There are misused or inappropriate cases or some overuse/underuse of connectives.	Several connectives might be misused or absent. Or there are frequent overuse of some simple connectives.	Connectives are often misused or absent.
Reader-writer interaction	Complexity of hedges and boosters	[F10]	There is rich variety of hedges and boosters. (with hedges more than boosters and in proper balance)	The hedges and boosters are used effectively, conveying a convincing and engaged stance.	There are frequent use of some simple words; occasionally with grammar mistakes. Or the hedge/booster words sometimes do not fit the context.	Reader-writer interaction features are sometimes used in writer's claims, but may be lack of convincing stance.	Reader-writer interaction features are rarely or not evident.

Appendix B: Writing prompts for a total of 18 essays

Year 1 Semester 1 Psychology course

Chloe 1, Lydia 1, and Rachel 1: Identify ONE past first-person experience related to education, and write a 1,000-word (±10%) essay analyzing the experience described, applying psychological concepts/principles to explain how the experience has occurred, and your behaviors, feelings or thoughts related to that experience.

Lucia and Wei: Identify one news report (commentaries and research articles are not allowed) within the 6 months preceding submission deadline, and write a 1,000-word (+/-10%) essay analyzing the incident or phenomenon described in the news clip, applying psychological concepts/principles to explain human behavior in educational settings. The news report should NOT be longer than 2 pages.

Year 1 Semester 2 Philosophy course (Chloe, Lydia, Rachel)

Choose an issue or topic that has been examined in the course that is especially relevant to your educational context (2000 words). Here, you will need to explain your educational context, be clear about the issue being addressed, provide rationale for your choice of issue, and explore how educators are approaching or could approach the issue.

Year 1 Semester 2 General Education course

Chloe: What are the main constraints and solutions to the provision of public open spaces in Hong Kong?

Lydia: Do innovation and entrepreneurship only happen in Start-up? Could 'elephant' dance nimble?

Lucia and Tiffany: 'Do you think religion plays a useful role in society nowadays'. Write 1,500~2,000 words, excluding references and other components of the ePortfolio.

Wei: 'To what extent can the concept of Confucianism, which were discussed in Professor Dennis Cheng's lectures, be applied to the people and society nowadays?'

Year 1 Semester 2 History course (Tiffany)

Examine one specific historical subject (e.g. a person, theme, development, event etc.) in relation to the specific contribution it has made to the origins of the Modern world. Write an essay between 1,500-1,800 words.

Year 2 Semester 2 General Education course

Rachel: Select a current issue about the conservation of living organisms (or their habitats) and write an essay about it. Your essay should explore the relationships between humans and the living world in the context of the selected issue. In discussing such relationships, you need to make reference to the different values that humans have towards the living world as described by Kellert (1996) and to other values as you see fit. You also need to discuss why conservation is important with respect to that particular issue and suggest possible ways to remedy the current situation and achieve any worthwhile goals.

Tiffany: Write a paper on a chosen natural disaster (that has not been covered by any group presentation in this course). The paper should include the causes of this natural disaster you select, its impact, and recovery strategies/measures to help affected people and areas to recover from this natural disaster.

Year 2 Semester 1 Intercultural communication course (Lucia)

Critically evaluate the values of your own culture (i.e. mainland China, HK, Nepal, Germany…etc). Write in 1500 words (plus or minus 150 words). When you write this essay, think about the theories, cultural patterns, different values and norms ...etc make reference to knowledge you learned from the course.

Year 2 Semester 1 SEN course essay (Wei)

Write an essay of not less than 2,200 words (in English) about a study of an issue or a case (empirical or non-empirical) related to teacher-parent relations or parent involvement in school education, making use of the theories, models, strategies and skills learnt. Make sure you support your essay with sound arguments based on relevant conceptual perspectives and

essential literature. Make use of your in class reflection exercises to articulate in the essay your critical reflection on home-school collaboration.

Appendix C: Coding scheme for interview protocol, writer log and open-ended survey items

Appendix C A systemic list of target influencing factors/sub-factors in L2 academic writing

Main factor	Sub-factor	Some details
I Discourse features	1 <topic building>	/topic analysis/, /focus clarification/, /thesis statement/
	2 <global coherence>	/introduction/, /supporting details/, /topic sentence/, /summary/, /conclusion/, /interface in-between/
	3 <local coherence>	/topic sentence supported by following details/, /sentence connections/, /reference nouns/, /idea flow/
	4 <logical connective>	/use of conjunctions and adverbial adjuncts/, /range of connectives/
	5 <reader-writer interaction>	/proper linguistic presentation of own attitude and opinion/, /concern on reader effect/
II Processing strategies	6 <research>	/task demand/ (writing purpose ect.), /length concern/, /search of related sources/ (width, depth, relevancy, ect), /academic reading/ (speed), /problem formulation/, /argument specification/, /overall research ability/
	7 <planning>	/idea generation/, /idea organization/, /degree of details in idea or organization planning/, /deliberating over the audience reaction/,, /sufficiency of research & planning/ (need of early research & planning, quality, quantity),
	8 <formulation>	/formulation efficiency/(length too long or short, speed), /integration of source materials/, /paraphrasing skills/, /calibration with other processing effects/，/consistency with planning/
	9 <revision>	/self-detection of problems/ （mainly grammar), /sufficiency of revision/(adequateness or degree of revision)
	10 <collaboration>	/co-creation/, /peer or teacher consultation/
	11 <evaluation>	/orchestration of the processing strategies/，/monitoring the whole composing processes/ (/distribution of time among strategies/, /time

		management skills/, /delayed beginning/)， /reflection of the overall composing process/
III Academic writing knowledge	12 <language skill>	/overall L2 proficiency/, /lexical complexity/, /grammar/, /sentence pattern/, /accuracy/, /fluency/, /conciseness/, /formality/, /other mechanics such as spelling/， , /sentence paraphrasing skills/， /machine translation/
	13 <genre knowledge>	/need of model essay/; /rhetorical conventions commonly practiced in target disciplinary communities/, /reference, citation & documentation format/, /Voice and identity construction (confident, objective, prudent, etc.)/, /anticipating and addressing audience's questions and needs/
	14 <disciplinary knowledge>	, /creative thinking/, /knowledge of research methodology/, /disciplinary background knowledge/ (width, depth, application, accuracy）
	15 <reasoning & argumentation>	/logic in argumentation/, /analytical reading and interpretation of published literature/, /critical synthesis of relevant literature/, /effective data or primary-text analysis /, /effective use and critique of individual studies/，/good thinking/ (clear, quick, systemic, flexible， deep)
IV Other	16 <overall writing proficiency>	/writing proficiency/, /L2 writing proficiency/, /influence of mother tongue/(Chinglish, influence of Chinese way of thinking)
	17 <attitude & emotions>	/careful/, /proactive/, /disengaged/, /lazy/， /dedicated/，/anxiety/, /ease of mind/
	18 <memory>	/memory/

Appendix D: Sample writing prompts for the 500-word essay

Instruction of 500-word Essay for *Integrated Business Challenges* (IBC course)

Assessment: 500 word essay (120 minutes)
Weighting: 20% of your overall mark
Task: You will each be given one of three different 'Cultural/Ethical Dilemmas' where your task will be to read the brief text and then research and apply the principles of problem solving to the dilemma. You have a week to prepare.

In your essay, you will be required to provide an analysis of both the cultural and ethical perspectives to the dilemma and required to demonstrate correct referencing which includes examples of in-text citations, direct quotations, paraphrasing and a reference list. (No referencing, both in-text and reference list, in your essay will result in a fail grade for this assessment. Include any problem-solving steps, such as a KT Analysis or Duncker as an appendix. This will not form part of your word count.)

You will be allowed to bring to class, a 1 page (A4 size) sheet of paper (1 side only) with hand written notes of your choice. This should include any quotes or references that you wish to include in your essay only. It cannot include any pre-writing essays or paragraphs in your hand-written notes. In addition, the notes will be collected with your essay papers.

Here is a brief suggestion (marking guide) as how to address the dilemma:

1. Identify the issue.
   Be clear about what the problem is.
   Remember that different people might have different views of what the issues are.

2. Understand everyone's interests. (4 marks)
   Interests are the needs that you want satisfied by any given solution.
   The best solution is the one that satisfies everyone's interests—if that was possible.

3. List the possible solutions (options) (4 marks)
   This is the time to do some brainstorming. There may be lots of room for creativity.

4. Evaluate the options. (supported by appendix) (3 marks)

What are the advantages/pluses and disadvantages/minuses?

Separate the evaluation of options from the selection of options.

5. Select an option or options. (2 marks)

   What's the best option, in the balance?

   Is there a way to 'bundle' a number of options together for a more satisfactory solution?

   Do you have a suggestion for dealing with any problem/issue unresolved? (1 mark)

The assessment will be under exam conditions. No copying. No talking. No use of the internet. No use of other electronic equipment of any sort.

Appendix E: The survey on students' writing strategy and knowledge

《综合商务挑战》课程 500-word essay 写作策略和写作知识现状调查

姓名：　　性别：　　班级：　　高考英语成绩：　　高考所在城市：

同学们好！ 本问卷基于你在本学期《综合商务挑战》课程 500-word essay 经历，从**写作策略运用**和**写作知识储备**两方面调查你在写作中遇到的具体困难和存在的典型问题。所选答案没有对错之分，请据个人真实情况作答，在对应数字上打勾（√）。你的选择将对中外合作办学学生**英文写作能力发展策略**研究提供重要借鉴。谢谢你的真诚反馈！

完全不符合	不符合	不太符合	基本符合	符合	完全符合
1	**2**	**3**	**4**	**5**	**6**

第一部分：写作策略调查

Academic writing strategies		完全不符合…… 完全符合
1	写作前，我已理解清楚**本写作任务的具体话题**/焦点问题。	1 2 3 4 5 6
2	写作前，我已理解清楚**题目中的写作要求**。	1 2 3 4 5 6
3	写作前，我已对相关阅读材料有充分理解。	1 2 3 4 5 6
4	我会把阅读材料和写作要求联系起来。	1 2 3 4 5 6
5	我已就写作话题，搜集相关资料，做好知识储备。	1 2 3 4 5 6
6	正式写作之前，我已对**文章结构**有清晰的认识。	1 2 3 4 5 6
7	我会根据老师要求和自己的写作进度，合理安排好写作时间。	1 2 3 4 5 6
8	正式写作之前，我已**形成一个具体要回答的问题**，来组织相关思路。	1 2 3 4 5 6
9	正式写作之前，我已思考要表达的**核心观点**（thesis statement）。	1 2 3 4 5 6
10	正式写作之前，我有构思要写的主要内容或观点（**main supporting points**）。	1 2 3 4 5 6
11	在课外提前完成相对完整的论文稿子的话，针对相关**词汇、句法或文章结构，**我会进一步检查。	1 2 3 4 5 6
12	写 Introduction 段落时，我清楚需要包含哪些主要内容点。	1 2 3 4 5 6
13	写 Concluding 部分时，我清楚需要包含哪些主要内容点。	1 2 3 4 5 6
14	写作中，我会尽力在每个主体段落的开头写出 topic sentence, 便于段落大意更易被理解。	1 2 3 4 5 6
15	具体到段落内写作时，我会留意段落内**句子间的连贯和衔接**。	1 2 3 4 5 6
16	写作中，我会思考如何**运用恰当的衔接词**把句子或段落有效连接起来。	1 2 3 4 5 6

17	在考场完成论文稿子后，我能够留出一些时间做回顾检查。	1 2 3 4 5 6
18	在考场完成论文稿子后，我尽可能检查相关词汇或句法表达。	1 2 3 4 5 6
19	在考场完成论文稿子后，针对查出的问题，我会尽力做修改。	1 2 3 4 5 6
20	在考场完成论文稿子后，我主要修改的是词汇或句法表达。	1 2 3 4 5 6
21	在考场完成论文稿子后，我会对照题目要求**再检查一下相关内容或结构**。	1 2 3 4 5 6
22	无论是在考前或者在考场上，稿子完成后，我会思考自己**是否达到了写作要求**。	1 2 3 4 5 6
23	本次写作任务完成后，我会**回顾反思整个过程**，以期待下次进行类似任务时做得更好。	1 2 3 4 5 6
24	本次写作任务完成后，我会总结本次写作中自己存在的问题或遇到的困难主要是哪些方面。	1 2 3 4 5 6

第二部分：写作知识。针对《综合商务挑战》500-word essay 写作，和与专业学习相关的其它写作活动任务时，以下 19 道题目旨在了解你的写作知识储备现状。请认真思考以下题目，选择对应数字表达**你在多大程度上同意**所述观点：

完全不符合	不符合	不太符合	基本符合	符合	完全符合
1	**2**	**3**	**4**	**5**	**6**

Academic writing knowledge		完全不符合……完全符合
1	我能够正确运用英语语言，包括语法、拼写和标点符号。	1 2 3 4 5 6
2	我有相关的专业词汇储备。	1 2 3 4 5 6
3	我能够区分**口语**和**书面语**的特点。	1 2 3 4 5 6
4	我能够用英语**简洁流畅地进行书面**表达。	1 2 3 4 5 6
5	我能够就语法准确性，词汇运用、拼写或标点符号方面，做**自评和修改**。	1 2 3 4 5 6
6	我熟悉参考文献的引用格式。	1 2 3 4 5 6
7	我明白如何**清晰恰当地表达自己的立场和观点**。	1 2 3 4 5 6
8	我知道要预判读者**在内容方面的阅读需求**。	1 2 3 4 5 6
9	我知道要预判读者在**结构或语言表达方面**的阅读需求。	1 2 3 4 5 6
10	我知道如何进行**文献检索**。	1 2 3 4 5 6
11	我有关于**论文写作需要用到的研究方法**的知识。	1 2 3 4 5 6

12	我有相关的专业知识背景。	1 2 3 4 5 6
13	我知道如何**对文献进行解读分析**。	1 2 3 4 5 6
14	我知道如何**整合他人的观点**。	1 2 3 4 5 6
15	我知道如何在写作中采用**数据或文本分析**。	1 2 3 4 5 6
16	我知道如何在写作中有效**运用、评论**现有研究。	1 2 3 4 5 6

第三部分：Open-ended questions（可用中文、英文或中英结合回答）

针对 **pre-writing preparation**—**while-writing drafting** —**after-writing revision** —**essay submission**，
回顾这一过程，

a. 我认为《综合商务挑战》500-word essay 写作**考察的主要能力或知识是（请列出 2~3 点）：**

b. **自己相对优势(strengths)主要是**（请列出 2~3 点）：

c. 自己写作方面**主要的困难和不足**（difficulties or problems），或**下次进行类似写作需要改进的**是（请列出 2~3 点）：

Appendix F: Overall percentage distribution of participants' own perception

Table 4.4 Overall percentage distribution of participants' own perception (total heads: 277, with 46 HPs, 122 MPs and 109 LP)

Category			Strengths (total: percentage)	Weaknesses (percentage)
I. Discourse features	1	<topic building>	NA	NA
	2	<global coherence>	NA	/conclusion/(1%), /topic sentence/(1%),
	3	<local coherence>	NA	NA
	4	<logical connective>	/use of conjunctions and adverbial adjuncts/(1%),	/use of conjunctions and adverbial adjuncts(1%)
	5	<reader-writer interaction>	NA	NA
II. Processing strategies	1	<research>	/search of related sources/(17%) , /task demand/(11%), /problem formulation/(6%), /argument specification/(3%), /academic reading/(2%)	/search of related sources/(6%) , /task demand/(4%), /academic reading/(4%) , /problem formulation/(4%), /argument specification/(1%)
	2	<planning>	/idea organization/(26%), /idea generation/(4%) /sufficiency of research & planning/(11%)	/idea organization/ (13%), /idea generation/(1%), /sufficiency of research & planning/(5%)
	3	<formulation>	/formulation efficiency/(3%)	/formulation efficiency/(5%)
	4	<revision>	/self-detection of problems/(3%), /sufficiency of revision/(1%)	/self-detection of problems/(1%), /sufficiency of revision/(1%)
	5	<collaboration>	/peer or teacher consultation/(5%), /co-creation/(3%),	NA

	6	<evaluation>	/orchestration of the processing strategies/(4%), /monitoring the whole composing processes/(1%)	/orchestration of the processing strategies/(5%), /monitoring the whole composing processes/(3%), /reflection of the overall composing process/(1%)
	1	<language skill>	/lexical complexity/(9%), /overall L2 proficiency/(7%), /grammar/(6%), /accuracy/(3%), /sentence pattern/(2%), /conciseness/(1%), /formality/(1%), /fluency/(1%),	/lexical complexity/(58%), /grammar/(30%), /sentence pattern/(13%), /accuracy/(5%), /overall L2 proficiency/(3%), /conciseness/(3%), /formality/(3%), /other mechanics such as spelling/(2%)，/fluency/(2%), /machine translation/(1%), /sentence paraphrasing skills/(1%)，
III. Academic writing knowledge	2	<genre knowledge>	/reference, citation & documentation format/(5%), /Voice and identity construction/(3%), /anticipating and addressing audience's questions and needs/(1%)	/reference, citation & documentation format/(5%)
	3	<disciplinary knowledge>	/disciplinary background knowledge/(11%), /creative thinking/(10%), /knowledge of research methodology/(1%)	/disciplinary background knowledge/(14%), /creative thinking/(8%), /knowledge of research methodology/(3%)
	4	<reasoning & argumentation>	/logic in argumentation/(18%), /analytical reading and interpretation of published literature/(12%), /critical synthesis of relevant literature/(8%), /good thinking/(7%), /effective use and critique of individual studies/(4%)，	/logic in argumentation/(8%), /analytical reading and interpretation of published literature/(8%), /effective use and critique of individual studies/(7%), /critical synthesis of relevant literature/(5%), /good thinking/(1%)

IV. Other	1	<overall writing proficiency>	/writing proficiency/(8%), /L2 writing proficiency/(4%),	/writing proficiency/(5%), /L2 writing proficiency/(4%), /influence of mother tongue/(3%)
	2	<attitude & emotions>	/careful/(6%), /proactive/(3%), /dedicated/(1%),	/careful/(1%)
	3	<memory>	/memory/(1%)	NA

Appendix G: Detailed percentage distribution of participants' own perception

Table 4.7 Detailed percent distribution of participants' own perception

(Total heads in terms of L2 proficiency: 277, with 46 HPs, 122 MPs and 109 LP; that of writing proficiency, 273, with 107 HPs, 120 MPs and 46 LPs)

Category			Strengths (L2 proficiency vs L2 writing level)		Weaknesses (L2 proficiency vs L2 writing level)	
I. Discourse features	1	<topic building>	H/M/L	NA	H/M/L	NA
	2	<global coherence>	HP	NA	HP	/topic sentence/(0 vs 2%), /supporting details/(0 vs 1%) , /conclusion/(0 vs 1%)
			MP	NA	MP	/topic sentence/(2% vs 0), /introduction/(0 vs 1%), /conclusion/(1% vs 1%)
			LP	/supporting details/(1% vs 2%)	LP	/conclusion/(2% vs 0), /introduction/(1% vs 0), /supporting details/(1% vs 0)
	3	<local coherence>	H/M/L	NA	H/M/L	NA
	4	<logical connective>	HP	/use of conjunctions and adverbial adjuncts/(2% vs 2%)	HP	/use of conjunctions and adverbial adjuncts(4% vs 1%)
			MP	/use of conjunctions and adverbial adjuncts/(2% vs 1%)	MP	/use of conjunctions and adverbial adjuncts(0 vs 1%)
			LP	/use of conjunctions and adverbial adjuncts/(1% vs 0)	LP	/use of conjunctions and adverbial adjuncts(1% vs 1%)

	5	<reader-writer interaction>	H/M/L	NA	H/M/L	NA
II. Processing strategies	1	<research>	HP	/task demand/(9% vs 11%), /search of related sources/(22% vs 15%) , /academic reading/(4% vs 2%) , /problem formulation/(4% vs 5%), /argument specification/(2% vs 2%)	HP	/task demand/(2% vs 5%), /search of related sources/(4% vs 6%) , /academic reading/(2% vs 5%), /problem formulation/(2% vs 6%), /argument specification/(0 vs 2%)
			MP	/task demand/(11% vs 14%), /search of related sources/(12% vs 22%), /academic reading/(2% vs 2%), /problem formulation/(7% vs 9%), /argument specification/(2% vs 4%)	MP	/task demand/(5% vs 3%), /search of related sources/(8% vs 3%), /academic reading/(4% vs 3%), /problem formulation/(4% vs 3%), /argument specification/(2% % 1), /overall research ability/(0 vs 1%)
			LP	/task demand/(12% vs 4%), /search of related sources/(21% vs 13%), /academic reading/(0 vs 2%), /problem formulation/(6% vs 2%), /argument specification/(3% vs 4% vs 0)	LP	/task demand/(5% vs 7%), /search of related sources/(4% vs 13%) , /academic reading/(4% vs 2%) , /problem formulation/(4% vs 0), argument specification/(2% vs 2%), /overall research ability/(1% vs 0)
	2	<planning>	HP	/idea organization/(35% vs 30%), /idea generation/(4% vs 2%), /sufficiency of research & planning/(7% vs 15%)	HP	/idea organization/ (20% vs 8%), /sufficiency of research & planning/(0 vs 5%)

		MP	/idea organization/(28% vs 26%), /idea generation/(6% vs 5%), /sufficiency of research & planning/(14% vs 7%)	MP	/idea organization/(9% vs %16), /idea generation/(2% vs 1%), /sufficiency of research & planning/(0 vs 4%)
		LP	/idea organization/(21% vs 20%), /idea generation/(1% vs 2%), /sufficiency of research & planning/(10% vs 13%)	LP	/idea organization/(14% vs 15%), /idea generation/(1% vs 4%), /sufficiency of research & planning/(0 vs 7%)
3	<formulation>	HP	/formulation efficiency/(2% vs 6%), /consistency with planning/(2% vs 0)	HP	/formulation efficiency/(4% vs 4%)
		MP	/formulation efficiency/(4% vs 1%), /consistency with planning/(0 vs 1%)	MP	/formulation efficiency/(4% vs 6%), /consistency with planning/(1% vs 1%)
		LP	/formulation efficiency/(2% vs 2%)	LP	/formulation efficiency/(6% vs 7%)
4	<revision>	HP	/self-detection of problems/(2% vs 3%), /sufficiency of revision/ (0 vs 2%)	HP	/self-detection of problems/(0 vs 2%), /sufficiency of revision/(0 vs 3%)
		MP	/self-detection of problems/(3% vs 2%), /sufficiency of revision/(1% vs 1%)	MP	/self-detection of problems/(0 vs 1%), /sufficiency of revision/(2% vs 1%)
		LP	/self-detection of problems/(2% vs 2%), /sufficiency of revision/(1% vs 0)	LP	/self-detection of problems/(4% vs 2%), /sufficiency of revision/(1% vs 0)
5	<collaboration>	HP	/peer or teacher consultation/(2% vs 4%), /co-creation/(0 vs 4%)	HP	/peer or teacher consultation/(2% vs 1%)
		MP	/peer or teacher consultation/(8% vs 6%), /co-creation/(4% vs 2%)	MP	/co-creation/(1% vs 1%)
		LP	/peer or teacher consultation/(3% vs 9%), /co-creation/(4% vs 4%)	LP	NA

			HP	/orchestration of the processing strategies/(0 vs 4%), /monitoring the whole composing processes/(0 vs 1%), /reflection of the overall composing process/(0 vs 1%)	HP	/orchestration of the processing strategies/(7% vs 6%), /monitoring the whole composing processes/(0 vs 5%)
	6	<evaluation>	MP	/orchestration of the processing strategies/(6% vs 2%), /monitoring the whole composing processes/(1% vs 1%)	MP	/orchestration of the processing strategies/(5% vs 5%), /monitoring the whole composing processes/(2% vs 2%), /reflection of the overall composing process/(2% vs 2%)
			LP	/orchestration of the processing strategies/(3% vs 9%), /monitoring the whole composing processes/(1% vs 0), /reflection of the overall composing process/(1% vs 0)	LP	/orchestration of the processing strategies/(5% vs 4%), /monitoring the whole composing processes/(6% vs 2%), ,/reflection of the overall composing process/(0 vs 2%)
III. Academic writing knowledge	1	<language skill>	HP	/overall L2 proficiency/(16% vs 7%), /grammar/(13% vs 6%), /lexical complexity/(11% vs 7%), /accuracy clarity precision/(0 vs 3%), /fluency/(0 vs 2%), /sentence pattern/(0 vs 3%), /formality/(0 vs 4%)	HP	/overall L2 proficiency/(2% vs 4%), /lexical complexity/(59% vs 57%), /grammar/(15% vs 33%), /sentence pattern/(17% vs 11%), /accuracy clarity precision/(9% vs 7%), /formality/(4% vs 3%), /conciseness/(2% vs 2%), /sentence paraphrasing skills/(2% vs 1%), /machine translation/(2% vs 2%), /other mechanics such as spelling/(0 vs 2%)

		MP	/overall L2 proficiency/(10% vs 8%), /grammar/(7% vs 7%), /lexical complexity/(11% vs 10%), /sentence pattern/(2% vs 2%), /accuracy clarity precision/(2% vs 3%), /fluency/(2% vs 0), /conciseness/(2% vs 2%), /formality/(1% vs 0)	MP	/overall L2 proficiency/(2% vs 3%), /lexical complexity/(56% vs 58%), /grammar/(35% vs 28%), /sentence pattern/(9% vs 14%), /accuracy/(6% vs 6%), /conciseness/(4% vs 5%), /other mechanics such as spelling/(3% vs 2%), /formality/(2% vs 3%), /fluency/(2% vs 3%), /machine translation/(2% vs 1%), /sentence paraphrasing skills/(1% vs 0)
		LP	/overall L2 proficiency/(2% vs 4%), /grammar/(3% vs 7%), /lexical complexity/(6% vs 11%), /accuracy clarity precision/(6% vs 2%), /fluency/(0 vs 2%), /formality/(3% vs 0), /sentence pattern/(2% vs 0), /conciseness/(1% vs 2%)	LP	/overall L2 proficiency/(6% vs 2%), /lexical complexity/(60% vs 61%), /grammar/(29% vs 26%), /sentence pattern/(16% vs 15%), /accuracy clarity precision/(4% vs 2%), /formality/(4% vs 4%), /fluency/(3% vs 0), /conciseness/(3% vs 2%), /other mechanics such as spelling/(2% vs 4%)
2	<genre knowledge>	HP	/reference, citation & documentation format/(9% vs 5%), /voice and identity construction/(2% vs 3%), /anticipating and addressing audience's questions and needs/(0 vs 1%)	HP	/reference, citation & documentation format/(9% vs 4%), /voice and identity construction/(0 vs 1%), /anticipating and addressing audience's questions and needs/(0 vs 1%)

		MP	/reference, citation & documentation format/(3% vs 6%), /voice and identity construction/(2% vs 3%), /anticipating and addressing audience's questions and needs/(0 vs 1%)	MP	/reference, citation & documentation format/(5% vs 6%), /voice and identity construction/(1% vs 0), /anticipating and addressing audience's questions and needs/(1% vs 0)
		LP	/reference, citation & documentation format/(6% vs 4%) , /voice and identity construction/(4% vs 2%), /anticipating and addressing audience's questions and needs/(2% vs 0)	LP	/reference, citation & documentation format/(4% vs 7%)
		HP	/creative thinking/(11% vs 8%), /disciplinary background knowledge/(7% vs 11%), /knowledge of research methodology/(2% vs 0)	HP	/creative thinking/(9% vs 9%), /disciplinary background knowledge/(13% vs 12%) /knowledge of research methodology/(2% vs 4%)
3	<disciplinary knowledge>	MP	/creative thinking/(6% vs 8%), /disciplinary background knowledge/(13% vs 11), /knowledge of research methodology/(0 vs 1%)	MP	*/creative thinking/(10% vs 7%), /disciplinary background knowledge/(18% vs 14%)*
		LP	/creative thinking/(14% vs 17%), /disciplinary background knowledge/(10% vs 11%), /knowledge of research methodology/(1% vs 0)	LP	/creative thinking/(6% vs 9%), /disciplinary background knowledge/(9% vs 17%), /knowledge of research methodology/(6% vs 2%)

		HP	/logic in argumentation/(22% vs 24%), /analytical reading and interpretation of published literature/(13% vs 14%), /critical synthesis of relevant literature/(2% vs 10%), /effective use and critique of individual studies/(2% vs 3%), /good thinking/(9% vs 9%)	HP	/logic in argumentation/(7% vs 10%), /effective use and critique of individual studies/(9% vs 5%), /critical synthesis of relevant literature/(7% vs 6%), /analytical reading and interpretation of published literature/(4% vs 8%)
4	<reasoning & argumentation>	MP	/logic in argumentation/(16% vs 13%), /analytical reading and interpretation of published literature/(11% vs 10%), /critical synthesis of relevant literature/(10% vs 7%), /effective use and critique of individual studies/(2% vs 3%), , /good thinking/(7% vs 7%)	MP	/logic in argumentation/(12% vs 6%), /effective use and critique of individual studies/(8% vs 9%), /analytical reading and interpretation of published literature/(7% vs 8%), /critical synthesis of relevant literature/(4% vs 5%), /good thinking/(2% vs 2%)
		LP	/logic in argumentation/(19% vs 17%), /analytical reading and interpretation of published literature/(13% vs 11%), /critical synthesis of relevant literature/(8% vs 7%), /effective use and critique of individual studies/(6% vs 9%), /good thinking/(6% vs 2%)	LP	/logic in argumentation/(5% vs 7%), /analytical reading and interpretation of published literature/(10% vs 4%), /critical synthesis of relevant literature/(6% vs 4%), /effective use and critique of individual studies/(5% vs 4%)

			HP	/writing proficiency/(13% vs 7%), /L2 writing proficiency/(4% vs 7%)	HP	/writing proficiency/(0 vs 7%), /L2 writing proficiency/(0 vs 4%), /influence of mother tongue/(0 vs 6%)
	1	<overall writing proficiency>	MP	/writing proficiency/(5% vs 7%), /L2 writing proficiency/(4% vs 3%)	MP	/writing proficiency/(5% vs 3%), /L2 writing proficiency/(5% vs 3%), /influence of mother tongue/(3% vs 1%)
			LP	/writing proficiency/(8% vs 11%), /L2 writing proficiency/(4% vs 2%)	LP	/writing proficiency/(6% vs 4%), /L2 writing proficiency/(4% vs 4%), /influence of mother tongue/(3% vs 0)
IV. Other			HP	/careful/(7% vs 7%), /proactive/(0 vs 5%), /dedicated/(0 vs 1%)	HP	/careful/(0 vs 1%)
	2	<attitude & emotions>	MP	/careful/(7% vs 7%), /proactive/(4% vs 2%), /dedicated/(1% vs 1%)	MP	/careful/(1% vs 1%), /disengaged/(1% vs 1%), /lazy/(1% vs 1)
			LP	/careful/(5% vs 2%), /proactive/(4% vs 4%), /dedicated/(1% vs 2%)	LP	/careful/(1% vs 0)
	3	<memory>	HP	/memory/(2% vs 1%)	HP	NA
			MP	/memory/(0 vs 1%)	MP	/memory/(0 vs 1%)
			LP	/memory/(1% vs 0)	LP	/memory/(1% vs 0)

Appendix H: Pre-, post-, and delayed post-test writing tasks

Pretest task

Living alone or not?

In recent times, many university students are making the decision to room alone. What are the causes of this? Does it have positive or negatives consequence?

Write an essay to discuss the cause and consequence of university students' preference in rooming alone. Give reasons for your answer and include any relevant examples from your own experience or knowledge.

You have about 40 minutes on this task. Write at least 250 words.

Reading material Living Alone in a Social World （Excerpts p. 80-82）
Olds, J., & Schwartz, R. S. (2009). The lonely American: Drifting apart in the twenty-first century. Beacon Press.
(Take notes on what you read, use them to organise information before writing, and cite information from sources accurately.)

Post-test task

Handwriting vs typing: is the pen still mightier than the keyboard?

Taking notes on laptops or computers rather than by pen and paper is increasingly common. It is worried that students' increasing reliance on keyboards for taking notes may impair their learning capacity.

Write an essay to discuss the cause and consequences of students' use of electronic products for note-taking. Give reasons for your answer and include any relevant examples from your own experience or knowledge.

You have about 40 minutes on this task. Write at least 250 words.

Reading material (excerpt of a research article)
Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. Psychological science, 25(6), 1159-1168.

Laptop note taking has been rapidly increasing across college campuses. The use of laptops in classrooms is controversial. Many professors believe that computers and the Internet serve as distractions, detracting from class discussion and student learning. On the contrary, students

often self-report a belief that laptops in class are beneficial. Even when students admit that laptops are a distraction, they believe the benefits outweigh the costs. Empirical research tends to support the professors' view, finding that students using laptops are not on task during lectures, show decreased academic performance, and are actually less satisfied with their education than their peers who do not use laptops in class.

We conducted three experiments to investigate whether taking notes on a laptop instead of writing longhand affects academic performance. On multiple college campuses, using both immediate and delayed testing across several content areas, we found that participants using laptops were more inclined to take verbatim notes than participants who wrote longhand, thus hurting learning. Moreover, we found that this pattern of results was resistant to a simple verbal intervention. Telling students not to take notes verbatim did not prevent this deleterious behavior. One might think that the harms of taking notes by laptop would be partially offset by the fact that verbatim transcription would leave a more complete record for external storage, and this would allow for better studying from those notes. However, we found the opposite—even when allowed to review notes after a week's delay, participants who had taken notes with laptops performed worse on tests of both factual content and conceptual understanding, compared to participants who had taken notes longhand.

We found no difference in performance on factual questions in the first two studies. In Study 3, it is unclear why longhand note-takers outperformed laptop note takers on factual questions, as this difference was not related to the relative lack of verbatim overlap in longhand notes. It may be that longhand note takers engage in more processing than laptop note takers, thus selecting more important information to include in their notes. This enables them to study this content more efficiently. It is worth noting that longhand note takers' advantage on memory of factual content is limited to conditions in which there was a delay between presentation and test. This may explain the discrepancy between our studies and previous research (Bui et al., 2013). The tasks they describe would also fall under our factual-question category, and we found no difference in performance on factual questions in immediate testing. For conceptual items, however, our findings strongly suggest the opposite conclusion. Additionally, whereas Bui et al. (2013) argue that verbatim notes are superior, they did not report the extent of verbatim overlap, merely the number of ‘idea units.’ Our findings are in agreement with theirs in that more notes (and therefore more ideas) led to better performance.

The studies we report here show that laptop use can negatively affect performance on educational assessments, even—or perhaps especially—when the computer is used for its intended function of easier note taking. Although more notes are beneficial, at least to a point, if the notes are taken randomly or by mindlessly transcribing content, the benefit disappears. Indeed, synthesising and summarising content rather than verbatim transcription can serve as a desirable difficulty toward improved educational outcomes (e.g., Diemand-Yauman, Oppenheimer, & Vaughan, 2011; Richland, Bjork, Finley, & Linn, 2005). For that reason, laptop use in classrooms should be viewed with a healthy dose of caution; despite their growing popularity, laptops may be doing more harm in classrooms than good.

Delayed post-test task

The cause and consequences of students' decision on an intended major in university study

Read the provided text and write an essay based on your reading. You have about 40 minutes on this task. Write at least 250 words.

Choosing a major is a milestone in a student's academic career. While some students know what they want to study before entering university, many undergraduates seek additional information before deciding on a major, such as parental encouragement, salary expectation, and job opportunities. What may the causes of your decision on a business major in university study? Does it have positive or negatives consequence?

Write an essay to discuss the cause and consequences of choosing a business major in university study. Give reasons for your answer and include any relevant examples from your own experience or knowledge.

You have about 40 minutes on this task. Write at least 250 words.

Reading material (excerpt of a research article)
Pritchard, R. E., Potter, G. C., & Saccucci, M. S. (2004). The selection of a business major: Elements influencing student choice and implications for outcomes assessment. *Journal of Education for Business*, *79*(3), 152-156.

Existing research indicates when students are satisfied with their academic major they are better motivated, have a greater sense of belonging, and achieve higher grade-point averages, than those who are undecided. While some students know what they want to study before entering university, many undergraduates seek additional information before deciding on a major. Studies show students pursue specific majors for a variety of reasons including of personal interest, parental encouragement and/or job-related beliefs about topics like salary expectations, job openings, and career growth.

Why do students choose different business majors? Differences in attitudes among business majors may reveal differences in work expectations. Student misperceptions about careers related to their particular major may lead to future job dissatisfaction. For instance, students who associate a marketing major with fascinating advertising jobs, a management information systems major with high starting salaries, or a finance major with purely analytical work, may in fact have little idea of the reality involved in those choices. They may be unprepared for the required hard work, long hours, tedious tasks, and great amounts of time involved in dealing with a variety of people.

Our study supports the view that many students pursue the business degree primarily for job opportunities and high pay following graduation. However, student interest in work related to a major was given as the top reason for choosing that major. These results strongly suggest that business students choose their major for other reasons besides job opportunities

and good pay. Student interest in the nature of a job and ‘ability to succeed’ are just as important.

We noted similarities among the different business majors. First, students seem to be internally motivated in their choice of major. That is, their top reasons are interest in the type of work, job opportunities, opportunities for self-employment, good match with student abilities, and projected earnings. On the other hand, the reputation of the major at school, perceived quality of instruction, amount and type of promotional information, and parents' and friends' influence exert only a very small impact on that choice.

We also noted differences among business majors. Compared with other majors, a higher percentage of finance majors ranked projected earnings as more important whereas general business and management majors placed greater emphasis on how their specialisation would help them run their own business. Also, unlike other business majors, MIS/CIS (management/computer information systems) majors were driven more by job prospects than by interest in the related work.

In addition, accounting, marketing, MIS/CIS, and double majors indicated that understanding of major and choice of major were important for career success. Interestingly, finance majors believed that choice of major was important for career success but that understanding of major was not; management majors believed that choice of major, understanding of business would have no obvious influence on their career success.

Compared with other business majors, those who chose accounting, marketing, MIS/CIS, or the double major seemed most interested in pursuing a directly related career (e.g., accounting, auditing, programmer/analyst) or finding immediate employment following graduation. The other business majors' career goals were less clearly defined; their choice included plans to pursue a vast array of career options. General business and management majors, for instance, stood out as being most interested in self-employment as a career objective.

How students choose a business major may reflect their response to the realities of the marketplace. For instance, the well-publicised hiring needs of information technology majors during the recent economic boom influenced major selections. The same was with reports which stated that stable, long-term careers within a single organisation are no longer common.

Appendix I: Comparison for group means, standard deviations and performance across time

Table 5.3 Comparison for group means, standard deviations and performance across time

Construct	Group	Pre-test	Post-test	Delayed Post-test	ANOVA		Pre vs Post		Pre vs Delayed Post	
		Mean(SD)	Mean(SD)	Mean(SD)	F	p	p	Eta Squared	p	Eta Squared
[DO]	EG	2.191(.573)	2.868(.499)	2.385(.583)	9.538	.000	**.000**	.293	.447	.029
	CG	2.032(.536)	2.388(.599)	2.573(.568)	5.624	.005	.083	.093	**.004**	.200
[F1]	EG	2.500(.752)	3.583(.584)	2.750(.766)	15.519	.000	**.000**	.403	.441	.028
	CG	2.521(.521)	3.146(.683)	3.083(.747)	6.582	.002	**.004**	.216	**.011**	.166
[F2]	EG	2.042(.793)	2.333(.928)	2.167(.654)	.804	.452	.421	.029	.851	.008
	CG	1.833(.637)	2.083(.868)	2.104(.531)	1.134	.328	.429	.027	.371	.053
[F3]	EG	1.771(.780)	2.125(1.076)	2.021(1.016)	.853	.853	.417	.036	.644	.019
	CG	1.625(.663)	1.521(.853)	2.250(.944)	5.438	.006	.901	.005	**.029**	.133
[F4]	EG	2.167(.776)	3.313(.689)	2.500(.752)	15.242	.000	**.000**	.389	.269	.047
	CG	1.980(.634)	2.583(.789)	2.854(.773)	8.906	.000	**.016**	.157	**.000**	.286
[F5]	EG	2.541(.896)	3.021(.759)	2.563(1.035)	2.155	.124	.165	.080	..996	.000
	CG	2.208(1.141)	2.646(1.089)	2.792(1.151)	1.741	.183	.376	.039	.180	.063
[F6]	EG	2.125(.537)	2.833(.602)	2.313(.882)	6.787	.002	**.002**	.287	.616	.017
	CG	2.021(.683)	2.354(.878)	2.354(.810)	1.403	.253	.321	.045	.321	.049

Appendix J: Approval Letter from University Human Research Ethics Committee

香港教育大學
The Education University of Hong Kong

26 August 2019

Ms WANG Yumin
Research Postgraduate Programmes
Graduate School

Dear Ms Wang,

Application for Ethical Review <Ref. no. 2018-2019-0386>

I am pleased to inform you that approval has been given by the Human Research Ethics Committee (HREC) for your research project:

Project title: Diagnostic Assessment of Discourse Competence in Chinese EFL Undergraduate Students' English Academic Writing

Ethical approval is granted for the project period from 2 September 2019 to 30 June 2020. If a project extension is applied for lasting more than 3 months, HREC should be contacted with information regarding the nature of and the reason for the extension. If any substantial changes have been made to the project, a new HREC application will be required.

Please note that you are responsible for informing the HREC in advance of any proposed substantive changes to the research proposal or procedures which may affect the validity of this ethical approval. You will receive separate notification should a fresh approval be required.

Thank you for your kind attention and we wish you well with your research.

Yours sincerely,

Patsy Chung (Ms)
Secretary
Human Research Ethics Committee

c.c. Professor CHOU Kee Lee, Chairperson, Human Research Ethics Committee

香港新界大埔露屏路十號
10 Lo Ping Road, Tai Po, New Territories, Hong Kong
T (852) 2948 8888 F (852) 2948 6000 www.eduhk.hk