The Effects of Anxiety, Cognitive-task Variability and Planning Type on Chinese Students' Argumentative Writing in English as a Foreign Language

TING XU

Supervisors

Professor Lawrence J. Zhang and Professor Janet S. Gaffney

A thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy (Ph.D.) in Education

(Applied Linguistics and TESOL)

The University of Auckland, Auckland

New Zealand

2020

Abstract

This study was motivated by the lack of research that has explored the synergistic effects of cognitive-task variability and planning type on writing in a Chinese English-as-a-foreign-language (EFL) context. Although an increasing number of studies have investigated the effect of cognitive task complexity, along either resource-directing or resource-dispersing dimensions, on EFL writing, the results for the two dimensions of cognitive task complexity have been far from conclusive. These may be due mainly to the manipulations of cognitive demands within the tasks not being validated, redundant language measures used, and learner factors overlooked.

Based on the empirical gaps, this study: (a) Investigated whether the designed-to-be more complex task is indeed more cognitively demanding, (b) explored the isolated and synergistic impact of cognitive-task variability and planning type on EFL argumentative writing, and (c) examined the role language anxiety played in the influence of cognitive task complexity on argumentative writing in a Chinese EFL context.

A mixed-methods approach, which included self-report questionnaires, writing tasks and interviews, was used. Ninety-one Chinese English major undergraduates were recruited to validate, independently, the manipulation of cognitive-task variability in Phase I of the main study using dual-task methods and self-ratings. In Phase II of the main study, another 201 English majors were recruited as volunteers and randomly assigned to three equal groups: a comparison group and two experimental groups, who experienced different planning conditions (i.e., pre-task planning and online planning). Participants in each group were required to complete the *Second Language Writing Anxiety Inventory* and two argumentative writing tasks at different cognitive levels. Quantitative and qualitative methods were adopted to examine the relationships among language anxiety, cognitive-task variability, planning type and students' writing performance.

The quantitative data were subjected to statistical analysis, such as *t*-tests and ANOVAs, and a thematic analysis was conducted on the qualitative data. Results of Phase I of the main study supported the efficacy of the cognitive-task variability manipulations; that is, the complex task version, as intended, was more cognitively demanding than the simple version. In response to the second research question, increases in cognitive-task variability enhanced accuracy, fluency, adequacy, and perhaps, syntactic complexity in the pre-task planning group, consistent with Robinson's Cognition Hypothesis (CH). Increasing cognitive-task variability led to trade-offs between syntactic complexity and lexical complexity, and between syntactic complexity and adequacy in the online planning group, and trade-offs between lexical complexity and accuracy, and between accuracy and adequacy in the comparison group, which supports Skehan's Limited Attentional Capacity Model (LACM). Furthermore, different planning types were found to ease the pressure on the central executive in working memory in different ways, leading to varying effects on learners' allocation of attention and EFL production. With regard to the third research question, writing anxiety was quantitatively and qualitatively found to interfere with the effects of cognitive task complexity on EFL writing performance.

Based on the findings, this study proposed a Model of Cognitive Task Complexity in EFL Writing Processes by combining the oral models of cognitive task complexity (i.e., CH and LACM) with Kellogg's writing model to make the theories more suitable in the EFL writing field. Other theoretical, methodological and practical implications were also discussed.

Key words: cognitive-task variability, pre-task planning, online planning, writing anxiety, EFL writing

Dedication

To my family

Acknowledgements

First, my deepest sincere gratitude goes to my dear supervisors: Professor Lawrence Jun Zhang and Professor Janet S. Gaffney. I am fortunate to have such an excellent supervision team that provides me with professional guidance and continuous support. Professor Lawrence Jun Zhang is like a lighthouse to my PhD journey, constantly reminding me of the direction of my research without being off course. He has always been there when I feel puzzled not only for academic support but also for life guidance. His continuous and insightful advice has allowed me to proceed smoothly with much free space. Professor Janet S. Gaffney has provided me with valuable and insightful advice on my study. Her continuous and meticulous feedback on my chapter and thesis drafts enlightened me about academic writing. She has always been tolerant of my “dropping-by” and glad to answer my questions and share opinions with me. Without my supervisors’ generous support and encouragement, this thesis would not have been possible. They are the greatest supervisors as well as incredible life mentors!

I would like to extend my sincere appreciation to the staff and students of the Chinese universities for their participation in this research and co-operation during the data collection.

I would also like to thank my PhD peers, who gave me valuable statistical advice on data processing, as well as useful feedback on my study. The friendship I gained during my PhD journey will always be a treasure in my life.

I am also extremely grateful to the University of Auckland and China Scholarship Council for the joint doctoral scholarship, which helps me to fully concentrate on my PhD study and complete the thesis smoothly.

Words cannot express how grateful I am to my beloved family: my parents, my grandparents, my uncle and aunt, my parents-in-law and so on. Thanks for their everlasting love, trust, encouragement and support in every step of my PhD journey. Special thanks go to my husband, Bo, as the shared journey was less arduous because of the mutual support, company, encouragement, and love.

Contents

List of Tables

List of Figures

List of Abbreviations

C/T	Clauses per T-unit
CAF	Complexity, Accuracy and Fluency
CET	College English Test
CG	Comparison group
CG-H	Comparison group-high anxiety (group)
CG-L	Comparison group-low anxiety (group)
CH	Cognition Hypothesis
CN/C	Complex nominals per clause
CP/C	Coordinate phrases per clause
DC/C	Dependent clauses per clause
EFC/C	The ratio of error-free clauses
EFL	English as a foreign language
EP100	The number of errors per 100 words
Err/T	The number of errors per T-unit
IPOA	Input, Processing and Output anxiety (scale)
L2	Second language
LACM	Limited Attentional Capacity Model
LD	Lexical density
MLC	Mean length of clause
MLT	Mean length of T-unit
MoE	Ministry of Education of the People’s Republic of China
OLP	Online planning
OLP-H	Online planning-high anxiety (group)
OLP-L	Online planning-low anxiety (group)

PTP	Pre-task planning
PTP-H	Pre-task planning-high anxiety (group)
PTP-L	Pre-task planning-low anxiety (group)
SLWAI	Second Language Writing Anxiety Inventory
T/S	T-units per sentence
TDMEQ	Task Difficulty and Mental Effort Questionnaire
TEM	Test for English Majors
TTR	Type-token Ratio
WT/√2W	The total number of different word types divided by the square root of two times the total number of words

Chapter 1 Introduction

1.1 Research Context

English language plays a pivotal role in globalization and has been rapidly achieving the status of the lingua franca in political, economic and educational domains (McKinley & Rose, 2018; Woodrow, 2011). In China, English is perceived as a necessary and crucial second or foreign language for everyone (Jin & Cortazzi, 2002), as success in English is key to advanced education, more employment opportunities, better professional and social development (Woodrow, 2011).

The English curricula for English majors and non-English majors are quite different at the tertiary stage. According to the Ministry of Education of the People's Republic of China (MoE, 2007), English majors are required to develop not only English competence in terms of reading, writing, listening and speaking, but also comprehensive knowledge about linguistics and literature. The English curriculum for non-English majors is designed to foster a comprehensive knowledge of English focusing primarily on grammar, vocabulary and reading, while speaking and writing are usually set up as optional courses (MoE, 2007). Accordingly, undergraduate English majors in China are required to complete 2,000-2,200 hours of English study (MoE, 2007), while non-English majors are required to complete 320 hours of English study during their four-year university studies. English majors, therefore, have longer exposure to English study than their non-English major counterparts.

There are two stages for English majors during the four-year programme: Stage I (the first two years) during which students are taught basic English skills in terms of reading, writing, speaking and listening, and Stage II (third and fourth years) when students are strongly encouraged to complement their English study with optional courses, such as culture, history, society, literature and linguistics (MoE, 2007). In addition, English majors in their second and fourth years are required to take the Test for English Majors (TEM), Band 4 or 8, which is administered by the National Advisory Committee for Foreign Language Teaching on behalf of the Higher Education Department, Ministry of Education, China (Jin & Fan, 2011). Non-English majors need to pass the College English Test (CET), Bands 4 or 6.

Writing is a compulsory component in the Chinese English language teaching curriculum. According to MoE (2007), English majors need to take English writing courses in the first two years of undergraduate study. In the writing course, learners are instructed, first on linguistic knowledge and, second on paragraph writing and, finally, genre-based writing skills and discourse structure knowledge, such as narration, exposition and argumentation (MoE, 2007). Among these writing genres, argumentation has been a popular and effective component of English assessment tools, such as the CET and TEM (Teng & Zhang, 2016, 2020), since this genre could evaluate students' writing achievements in terms of linguistic competence, critical thinking and articulation of ideas (Hirose, 2003). Additionally, the TEM tests that English majors are required to complete are designed to evaluate their English proficiency in reading, writing, listening and translation. Thus, in the examination-based settings, writing not only plays an important role in English teaching but also exerts a great deal of pressure on second language (L2) writers.

Writing, a core skill in China's English language teaching, has always been the most challenging for Chinese students, and an aspect in which they are the weakest (Guo & Qin, 2010; Huang & Zhang, 2020). The traditional teaching approaches may exacerbate Chinese students' challenges with English writing. For example, teachers used to be mainly concerned with grammatical accuracy in writing exam performance, and focus primarily on accuracy in their written work when giving feedback (Kuiken & Vedder, 2019; Lee, 2008, 2009, 2011; Mao & Crosthwaite, 2019; Trebits, 2016). Although this probably agitates students, learners' anxiety, which can obstruct the success of language learning (Horwitz, 2001), has often been overlooked in the teaching process (Guo & Qin, 2010). Writing anxiety, one of the most important affective factors, was proposed by Daly and Miller (1975b) as "a situation and subject-specific individual difference that is concerned with a person's general tendency to approach or avoid writing" (p. 242). Guo and Qin (2010) reported that Chinese university students experienced L2 writing anxiety to a medium degree, which has a significant negative correlation with their writing performance. Affective factors have been ignored by teachers and learners, given that product-oriented approaches dominated China's language teaching instruction (Guo & Qin, 2010). Under these circumstances, MoE (2007) has called for greater attention to students' affective factors.

In the EFL teaching context, reforms in pedagogies and curricula of writing teaching have been conducted in recent decades in China. As an important methodological innovation in EFL teaching, task-based instructional approach, in which learners perform a series of tasks using their own linguistic resources, has greatly influenced syllabus design, teaching methodology and material development. Implementing the task-based language learning and teaching pedagogy facilitates, to some extent, the transformation of the writing-teaching process from a teacher- to a student-centered teaching method. Task-based L2 writing also plays an important role in L2 learners' interlanguage development "by providing opportunities for testing hypotheses about language and writing, and for processing language through form-meaning mapping and language-related episodes arising during L2 writing processes, and also via post-task feedback provided on the learners' written text" (Rahimi, 2016, p. 3). More empirical research on task-based EFL writing in a Chinese context is needed, given the relevant studies in this field in China are not systematic and far and few between (Liu, Wang, & Zhang, 2017).

1.2 Cognitive Task Complexity

In recent decades, since the introduction of task-based language pedagogy in the 1980s, there has been an increasing number of publications related to task-based language learning, teaching, and assessment (for example, Bygate, 2015; Ellis, Skehan, Li, Shintani, & Lambert, 2019; Skehan, 2018; Wen & Ahmadian, 2019). Researchers have reached a consensus that tasks hold a central place in second language acquisition (SLA) research and language pedagogy. Given the advocacy for task-based learning, attention has turned to the grading and sequencing criteria for pedagogic tasks (Ellis, 2003; Ellis et al., 2019; Long & Crookes, 1992; Robinson, 2001a, 2015; Skehan, 1996a, 1998, 2003, 2014b, 2018).

In task-based approaches to pedagogy and syllabus design, such as those of Long (1996), and Skehan (1996a, 1998, 2003, 2018), pedagogic tasks are sequenced not on the basis of their linguistic content but rather on various notions of task complexity. The concept of task complexity stems from the need to establish criteria for sequencing tasks in a principled way to foster interlanguage development. Research into the effects of task complexity on EFL learning is "an area of great consequence for the development of theories of instructed SLA and for pedagogic decisions about grading and sequencing tasks for learners" (Robinson, 2001a, pp. 316-317).

The current SLA theories of task complexity, such as Skehan's Limited Attentional Capacity Model (LACM) and Robinson's Cognition Hypothesis (CH), have identified a number of dimensions and variables by which task complexity is determined. Skehan (1996a, 1998, 2018) distinguished three factors contributing to the complexity of tasks: Code complexity (related to the linguistic demands of a task); cognitive complexity (concerning the processing demands of a task and the availability of task-related schematic knowledge); and communicative stress (related to time pressure, modality, scale and participant variables). Robinson (Baralt, Gilabert, & Robinson, 2014a; Robinson, 2001b, 2011a; Robinson & Gilabert, 2019) differentiated task complexity from task difficulty and task conditions. Task complexity can be manipulated along two dimensions: resource-directing dimensions that place cognitive/conceptual demands on participants and resource-dispersing dimensions that impose performative/procedural demands on learners. These dimensions affect learners' attentional resource allocation during EFL task performance differentially.

Skehan's and Robinson's models also address the issue of how manipulation of task factors affects EFL learning, performance and development. A large part of existing empirical research has set out to test the predictions for speaking (Foster & Skehan, 2013; Gilabert, 2007; Michel, Révész, Shi, & Li, 2019; Skehan, 2014a; Skehan & Shum, 2014; Wang & Skehan, 2014) and writing (Abrams, 2019; Ishikawa, 2007; Kang & Lee, 2019; Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2007a, 2008; Rahimi, 2019; Rahimi & Zhang, 2019; Révész, Kourtali, & Mazgutova, 2017). Research findings on the effects of cognitive task complexity on EFL writing performance, however, are far from conclusive, which signals the need for sustained investigation in this area.

1.2.1 Definitions of key term

Task complexity is defined by Robinson (2001b) as "the result of the attentional, memory, reasoning and other information processing demands imposed by the structure of the task on the language learner" (p. 29). "*Cognitive*" is added before task complexity to emphasize the interaction between the task and learners' cognition, as this interaction is involved in this study. As ***cognitive task complexity*** is a multi-dimensional construct, the definition "the inherent cognitive demands of a task that are realized in interaction with learner characteristics" (Sasayama, 2016, p. 232) was adopted in this study.

Robinson's terminology for the two dimensions of cognitive task complexity, *resource-directing* dimensions and *resource-dispersing* dimensions were applied in the current research. ***Resource-directing*** variables (e.g. contained elements, spatial location, causal relationships, and intentionality) distinguish task characteristics "on the basis of the concepts that the task requires to be expressed and understood" (Robinson, 2007a, p. 17). ***Resource-dispersing*** variables (e.g., planning, and familiar knowledge) "make increased performative/procedural demands on participants' attentional and memory resources, but do not direct them to any aspect of the linguistic system which can be of communicative value in performing a task" (Robinson, 2007a, p. 18). Based on a critical examination of the literature (see Chapter 2), *task complexity* has been short for "task complexity along *resource-directing* dimensions" in a number of studies (Cho, 2015; Frear & Bitchener, 2015; Kang & Lee, 2019; Kuiken & Vedder, 2007b, 2008, 2012; Rahimi & Zhang, 2018, 2019). These two constructs, however, are not exactly the same. Task complexity, according to Robinson, involves two dimensions: *resource-directing* dimensions and *resource-dispersing* dimensions. To avoid ambiguity and increase precision, the researcher uses the term ***cognitive-task variability*** to replace the *resource-directing* dimensions of cognitive task complexity. *Cognitive-task variability* is manipulated in this study by varying numbers of argument elements and reasoning demands. *Increasing cognitive-task variability*, particularly in the current study, refers to increasing cognitive task complexity along the resource-directing dimension by involving more argument elements and greater reasoning demands.

With the *resource-dispersing* variables, task complexity along *resource-dispersing* dimensions is manipulated by providing different **planning types** (i.e., pre-task planning and online planning). According to Ellis (2005), **pre-task planning** (PTP) and **online planning** (OLP), also called within-task planning, are "distinguished simply in terms of when the planning takes place—either before the task is performed or during its performance" (p. 3). The PTP condition is operationalized as 15 minutes of pre-task planning and 25 minutes to write. Under the OLP condition, participants are instructed to write immediately for 40 minutes and to plan during their writing (see more details in Section 3.6.3).

Cognitive task complexity can be manipulated by varying levels of cognitive demands, resulting in variation in learners' performance (i.e., complexity, accuracy, and fluency). The **complexity**, **accuracy** and **fluency** (CAF) of learners' performance can be seen as constituting a learner's language proficiency. Notional definitions of the three aspects, given by Skehan and Foster (1999), are shown in Table 1.1, although the operational definitions have varied in different studies (see literature review in Section 2.8).

Table 1.1 Notional Definitions of CAF (Skehan & Foster, 1999, pp. 96-97)

Aspect	Definition
Complexity	The capacity to use more advanced language, with the possibility that such language may not be controlled so effectively. This may also involve a greater willingness to take risks, and use fewer controlled language subsystems. This area is also taken to correlate with a greater likelihood of restructuring, that is, change and development in the inter-language system.
Accuracy	The ability to avoid error in performance, possibly reflecting higher levels of control in the language as well as a conservative orientation, that is, avoidance of challenging structures that might provoke error.
Fluency	The capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems.

1.3 Statement of Problems

The brief introduction and detailed literature review in Chapter 2 highlight the following gaps:

Firstly, previous empirical studies of effects of cognitive-task variability, or effects of planning on EFL writing, yielded mixed findings, which indicates the need for further investigations in this area. Additionally, little research has been conducted on the synergistic effects of cognitive-task variability and planning type on EFL writing.

Secondly, little research to date, has provided independent evidence in the field of EFL writing to triangulate that the designed-to-be complex task is more cognitively demanding than the simple version. Validation of the manipulations of cognitive-task variability, a key variable in cognitive task complexity, is essential (Baralt et al., 2014a; Révész, Michel, & Gilabert, 2016; Sasayama, 2016).

Thirdly, most prior studies typically used one or two index(es) of CAF to capture the effects of cognitive task complexity on L2 writing. Adequacy, an independent measure from CAF for pragmatic performance, was seldom used to tap the influence of cognitive task complexity. Multi-dimensional CAF measures as well as measures of adequacy are warranted to capture a multi-dimensional and consistent picture of EFL learners' writing performance (Housen, De Clercq, Kuiken, & Vedder, 2019; Kuiken, Vedder, Housen, & De Clercq, 2019; Norris & Ortega, 2009; Pallotti, 2009; Rahimi, 2019; Yang, Lu, & Weigle, 2015).

Finally, language anxiety, an affective factor that obstructs the success of language learning, may interfere with the effects of cognitive task complexity on EFL performance (Robinson, 2011a). The inclusion of L2 writing anxiety, as called for by researchers (Kuiken & Vedder, 2008; Rahimi & Zhang, 2019; Révész, 2011; Trebits, 2016) and the MoE (2007), is warranted in studies of cognitive task complexity.

1.4 Objectives and Research Questions

This study aims to: (a) Validate the cognitive-task variability manipulated within the two argumentative tasks; (b) explore the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing; and (c) examine the influence of writing anxiety on the effects that cognitive task complexity has on EFL writing in a Chinese context. The study is guided by the following research questions:

1. ***What are the relationships between presumed cognitive-task variability and cognitive load? If relationships exist,***

(a) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by self-perceived task difficulty and mental effort self-assessment?

(b) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by dual-task method?

2. ***What are the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing?***

(a) What is the potential effect of cognitive-task variability manipulated by varying the number of argument elements and reasoning demands on Chinese EFL students' argumentative writing?

(b) What is the potential effect of planning type (i.e. pre-task planning and online planning) on Chinese EFL students' argumentative writing?

(c) What are the potential synergetic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing?

3. ***What are the relationships among students' perceived writing anxiety, cognitive-task variability, planning type and EFL writing production?***

(a) What is the level of second language writing anxiety reported by university students in EFL writing?

(b) What are the potential effects of second language writing anxiety, cognitive-task variability and planning type on students' EFL writing?

(c) What are students' perceptions of second language writing anxiety related to cognitive-task variability?

(d) What are students' perceptions of second language writing anxiety related to planning type?

1.5 Significance of the Study

1.5.1 Theoretical contributions

This research builds on prior research and provides empirical evidence to clarify the dispute between CH and LACM on the effects of cognitive task complexity on EFL writing, because results of the existing studies into cognitive task complexity in the writing field are inconclusive.

The present study extends the understanding of cognitive task complexity in the field of EFL writing. No substantial attention has been given to the synergistic effects of the two dimensions in cognitive task complexity, although the two dimensions interact in students' writing performance in the real-world (Kormos, 2011). This study contributes to the enrichment of empirical evidence and theory on cognitive task complexity, by leading to a deeper understanding of the synergistic effects of cognitive-task variability and planning type on EFL writing. Additionally, while learner factors have been highlighted in Robinson's CH, very few studies have paid attention to such variables. This study explores how anxiety, one of the most important affective factors, interferes with cognitive task complexity to influence EFL writing, deepening understanding of the role of language anxiety plays in cognitive task complexity.

1.5.2 Methodological contributions

The present study contributes to innovations in efficiently measuring and independently validating the cognitive-task variability of EFL writing by adopting diverse methods from cognitive psychology for triangulation, such as dual-task method and self-rating. As for validation techniques, this study is an initial effort to use the dual-task technique to measure the cognitive demands imposed by writing tasks in the field of task-based research. In addition, this research is expected to capture a comprehensive picture of how EFL performance is influenced by cognitive task complexity by using multi-dimensional linguistic measures and adequacy. Finally, the mixed-method design extends existing research to investigate, and provide an in-depth understanding of the role of writing anxiety in cognitive task complexity.

1.5.3 Pedagogical implications

From a pedagogical perspective, this study is anticipated to provide additional empirical evidence to support the task-based instructional approach in order to enhance EFL writing teaching and learning. The inclusion of learner affective factors (i.e., writing anxiety), different planning types and validation of cognitive-task variability manipulation may broaden insights into task design and sequencing for pedagogy and assessment, as well as into the nature of EFL learners' attentional resources.

1.6 Outline of the Thesis

This thesis consists of seven chapters. Chapter One provides a general introduction to the study regarding research context, key terms, existing problems, objectives and research questions as well as the significance of the study. Chapter Two overviews the theoretical framework related to this research: Skehan's LACM, Robinson's CH and Kellogg's writing model. This chapter also provides a systematic literature review on the effects of cognitive task complexity (along two dimensions) on EFL writing, and the three possible main reasons for the inconclusive results of cognitive task complexity (i.e., validation of cognitive-task variability, measurements of EFL writing performance, and role of L2 writing anxiety). Chapter Three describes the methodology for the pilot study and for each stage of the main study. Chapter Four and Chapter Five respectively present the findings in response to the three sets of research questions. Chapter Six discusses the results reported in the preceding chapters. Chapter Seven concludes the endeavours by summarising research findings as a whole and by offering theoretical and practical implications. Limitations and suggestions for future research are also presented in this chapter.

Chapter 2 Theoretical Frameworks and Literature Review

This chapter starts with an introduction to the theoretical rationale for the research design from Section 2.1 to 2.3. Robinson's Cognition Hypothesis (CH) and Skehan's Limited Attentional Capacity Model (LACM), and Kellogg's Writing Model are presented, and the connections between cognitive task complexity and Kellogg's Model in the field of EFL writing are introduced. Following the theoretical framework, previous studies on the effects of cognitive task complexity are reviewed separately and synergistically along the two dimensions from Section 2.4 to Section 2.6. Three possible reasons for the divergent results in this field are discussed in Section 2.7, 2.8 and 2.9. Finally, the gaps identified in the field of task-based writing research are highlighted, followed by a summary of the chapter.

2.1 Task Complexity

Skehan (Skehan, 1996a, 1998, 2009, 2015, 2018; Skehan & Foster, 1999) and Robinson (Baralt, Gilabert, & Robinson, 2014b; Robinson, 2001b, 2005, 2011a, 2011b; Robinson & Gilabert, 2019), based on prior attempts at task sequencing in a syllabus, have respectively proposed two models of task complexity, LACM and CH, to identify a series of task design factors that can be manipulated to achieve varying levels of complexity. These two models also have attempted to address the question of how tasks with different levels of task complexity place different cognitive demands on learners, resulting in variation in learners' language performance.

2.1.1 The Limited Attentional Capacity Model

The LACM, which claims humans have a limited and single-resource attentional capacity (Skehan, 1996a, 1998; Skehan & Foster, 1999), has been developed based on the theories on working memory (Carter, 1998; Gathercole & Baddeley, 1993). The basic assumption of LACM is that "humans have a limited information processing capacity and must therefore prioritise where they allocate their attention" (Skehan, 2013, p. 189). As learners' attentional limits are reached on more demanding tasks, trade-off effects will occur among three aspects of language production: CAF (Skehan, 1996a; Skehan & Foster, 1999, 2001). In other words, learners can only attend to one aspect of language production (complexity, accuracy, or fluency) when completing a cognitively taxing task at a given time. According to Skehan (1996a), a focus on accuracy may make learners less willing to take risks, relying on less elaborate interlanguage systems and, at the same time, decrease fluency. A focus on complexity (i.e., the process of restructuring) "increases the chances that new forms will be incorporated into interlanguage systems, promotes risk-taking, and requires attention being devoted to the new forms of language" at the expense of accuracy and fluency (Skehan, 1996a, p. 50). A focus on fluency contributes to "language being produced more quickly to an emphasis on accessibility and with low priority being attached to getting language right, or to the use of new forms" (Skehan, 1996a, p. 50). Based on VanPatten's (1990, 1996) results, Skehan and Foster argued that there might be competition for limited working memory between a high level of cognitive processing and the local linguistic form. More specifically, if a task demands an allocation of working memory to content, less attentional resources will be available for language forms, leading to poor language performance in complexity and accuracy, which is true especially for EFL learners.

Skehan (1996a, 1998) has suggested three principles to analyse cognitive complexity of tasks: code complexity, cognitive complexity, and communicative stress. As shown in Table 2.1, code complexity is related to the linguistic demands of a task. A complex task is likely to require more advanced structures or greater densities of advanced structures. Cognitive complexity, including two areas of processing and familiarity, is concerned with the content or meaning of task performance, which relates to the conceptualization stage of Levelt's model (Skehan, 1996a). Processing refers to "the amount of on-line computation that is required while doing a task", and is closely related to "the extent to which the learner has to actively think through task content" (Skehan, 1996a, p. 52). A complex task is likely to require more online processing of the content than a simple task. The greater the processing demand, the higher the task complexity. In contrast, familiarity refers to the extent to which participants possess, and can readily use task-related schematic knowledge or information to complete the task at hand. The greater the familiarity, the lower the task complexity. Communicative stress refers to performance conditions, which result from differences in time pressure, modality, scale, and participant variables.

Table 2.1 Skehan's Model of Task Complexity in LACM

Code complexity	Cognitive complexity	Communicative stress
Vocabulary load and variety Redundancy and density	Cognitive familiarity • Familiarity of topic • Familiarity of discourse genre • Familiarity of task Cognitive processing • Information organization • Amount of computation • Clarity and sufficiency of information given • Information type	Time pressure scale Number of participants Length of texts used Modality Stakes Opportunity for control

These three dimensions influence the cognitive task complexity and consequently the way a learners' attention is allocated and how performance is affected during a task (Skehan & Foster, 2001). As a result, the three-way variables impact task difficulty. Task difficulty is related to the amount of attention the task requires of the learners. According to Skehan and Foster (2001), "difficult tasks require more attention than easy tasks" (p. 196).

Increases in cognitive task complexity lead to trade-off effects on language performance, due to the limited information processing capacity. Skehan (1998) claimed, therefore, that tasks should be designed and sequenced in a balanced manner to minimize the negative effects of the increased cognitive demands. Skehan further suggested that sequencing tasks from less to more difficulty is ideal, which does not push learners to allocate all attentional resources to conceptualizing the complex ideas at the beginning, but allows them to strategically devote their attention to language forms, promoting balanced language development in terms of CAF.

Skehan (2009) emphasized that the "natural" tension is between accuracy and complexity when learners' attentional limits are reached. The simultaneous enhancements in complexity and accuracy, as found in a few cases in previous research (Foster & Skehan, 1999, 2013; Tavakoli & Foster, 2008; Tavakoli & Skehan, 2005) are the result of the processing conditions, such as pre-task planning, post-task activities, and the joint operation of separate task factors, rather than increased task difficulty (as suggested by the Cognition Hypothesis).

Foster and Skehan (1999) explored the effects of three types of planning (i.e., solitary, group-based, and teacher-led planning) on EFL learners' oral performance. The results showed that teacher-led planning yielded joint increases in complexity and accuracy, possibly because the teacher-led planning helped combine the strengths of planning-as-rehearsal (that led to greater accuracy) and planning-as-complexification (that led to greater complexity), and ease attentional limitations. Foster and Skehan (2013), in examining the effects of post-task activities on EFL learners' oral performance in a decision-making task, found that participants who engaged in a post-task activity of transcribing their own utterances produced more complex and more accurate language than those with no transcription as a post-task activity. These findings echo Skehan's (1996a) claim that task implementation, such as pre-task, during and post-task activities, can help ease and free learners' limited attentional resources for the actual language during the task, and result in more complex, accurate and fluent language performance.

Moreover, a series of studies (Tavakoli & Foster, 2008; Tavakoli & Skehan, 2005) investigated the effects of task structure on narrative retelling performance and found that structured tasks led to greater accuracy, and that tasks, which required the integration of foreground and background information, assisted complexity performance. The results also showed that the task, supported by both structured feature and information integration, contributed to increased accuracy and complexity. Both structure and information manipulation in these studies were task-design features with the former assisting accuracy and the latter complexity; the influence of each could be additive and complementary. It seems to be the case that "under supportive conditions the constraints of limited attentional capacity can be overcome to some extent" (Skehan, 2009, p. 522).

To sum up, to enhance the pre-targeted dimension of learners' EFL production, assigned tasks should be manipulated at an appropriate level in terms of task difficulty (task types, task characteristics) and implementation.

Skehan (2009, 2015) used the Levelt's model of L1 speaking in relation to the L2 oral performance and suggested that the task-based performance results can be linked to this model. Complexification, pressuring, easing, and focusing as impacts were discussed by Skehan, relative to different stages of the Levelt model: Conceptualization, Lemma Retrieval Formulation, and Syntactic Encoding Formulation.

Table 2.2 The Levelt Model Linked to Influences on L2 Performance (Skehan, 2009, p. 525)

Complexifying/Pressuring		Easing/Focusing
• Planning: extending • More complex cognitive operations • Abstract, dynamic information • Greater quantity of information	**Conceptualizer**	• Concrete, static information • Less information • Less complex cognitive operations
• Need for less frequent lexis • Non-negotiability of task	**Formulator: Lemma Retrieval**	• Planning: organizing ideas • Dialogic
• Time pressure • Heavy input presence • Monologic	**Formulator: Syntactic Encoding**	• Planning: rehearsing • Structured tasks • Dialogic • Post-task condition

As shown in Table 2.2, in the left column are factors that augment the difficulty of a task. Factors falling into the Conceptualizer stage may force speakers to develop a message that requires more cognitive demands, pushing for greater language complexity. At the second stage, Lemma Retrieval Formulator, EFL learners need to exert great effort to ease their online processing problems and access more difficult lemmas during online speech production. Another set of pressuring impacts associated with Formulation stage of syntactic encoding, address the online pressures that speakers have to handle, including time pressure and the amount of input. Time pressure, as an example, is "an index of the amount of time the speaker has to access material, to build syntactic frames, and to regroup as necessary" (Skehan, 2009, p. 526).

By contrast, in the right-hand column are factors that ease either task difficulty or alternatively direct attention to a particular area. Firstly, the quantity and nature of information inherent in tasks may influence Conceptualization. For example, tasks with concrete and static information may ease cognitive demands on Conceptualizer operations. Secondly, easing factors in Formulator stages mainly function as freeing attention for lemma retrieval and syntactic building. For example, pre-task planning can assist learners to identify ideas, prime lexical elements, and prepare syntactic frames, thereby promoting smooth access to lemmas and syntactic structures during online processing.

To sum up, “complexification links mainly to the Conceptualization stage, and then structural and lexical complexity”, while “pressing, easing and focusing are more relevant for the Formulator, and then, accuracy and fluency” (Skehan, 2009, p. 527). The connection between task complexity and Levelt’s model can help understand how different task features and implementation conditions vary the amount of attention available and the focus of learners’ attentional resources when a message is expressed during speaking.

2.1.2 The Cognition Hypothesis

Robinson (2001b, 2005, 2011a, 2011b) proposed the task complexity and sequencing models for task-based EFL pedagogy, known as the Cognition Hypothesis, (also named Triadic Componential Framework), by incorporating Schmidt’s (2001) information-processing theories, Long’s (1996) interactionist explanations of L2 task effects and Wickens’ (1989, 1992) model of dual task performance. Contrary to LACM, CH holds that dimensions of cognitive task complexity belong to different attentional resource pools, and increases in cognitive demands of tasks may direct learners’ attentional resources to language form so that input may be processed more deeply and elaborately.

Robinson (2001b, 2005, 2011a, 2011b) claimed that cognitive complexity criteria can be operationalised across three dimensions that interact to influence task performance and learning: task complexity, task conditions, and task difficulty. The relationships of these dimensions in the Triadic Componential Framework are presented in Table 2.3.

Table 2.3 A Triad of Task Complexity, Task Condition and Task Difficulty Factors (Robinson, 2005, p. 5)

Task complexity (cognitive factors)	*Task conditions* (interactional factors)	*Task difficulty* (learner factors)
(a) resource-directing	(a) participation variables	(a) affective variables
e.g., ±few elements	e.g., open/closed	e.g., motivation
±Here-and-Now	one-way/two-way	anxiety
±no reasoning demands	convergent/divergent	confidence
(b) resource-dispersing	(b) participant variables	(b) ability variables
e.g., ±planning	e.g., same/different gender	e.g., working memory
±single task	familiar/unfamiliar	intelligence
±prior knowledge	power/solidarity	aptitude

Robinson stressed that task complexity should be distinguished from task conditions and task difficulty first. Task complexity refers to the characteristics that contribute to the intrinsic cognitive demands tasks make on learners. There are two types of cognitive task features that comprise task complexity: resource-directing and resource-dispersing variables. The resource-directing dimensions describe task characteristics "on the basis of the concepts that the task requires to be expressed and understood" (Robinson, 2007a, p. 17). The proposed resource-directing variables include whether the task requires reference to here and now, past or present events, few or many elements, and more or fewer reasoning demands. These variables place cognitive/conceptual demands on participants, and can direct learners' attention and effort during conceptualization to particular aspects of the language code system, thus facilitating "the development and acquisition of new L2 form-concept mappings" (Robinson, 2007a, p. 18). By contrast, resource-dispersing task characteristics impose performative and procedural demands on learners' cognitive resources and facilitate learners' automatic access to "an already established interlanguage system" (Robinson, 2007a, p. 18). The proposed resource-dispersing variables include: planning time, prior knowledge and the number of tasks to be completed.

According to Robinson (2001b, 2005, 2007a, 2011a), cognitive task complexity can be operationalized in a predictable way in these two dimensions, and the effects of task complexity along the two dimensions should be very different. Complexity, on the one hand, can be manipulated by varying resource-directing task variables: For example, by increasing or decreasing the number of elements involved in the task; by increasing or decreasing the reasoning demands of the tasks. Increasing task complexity along resource-directing dimensions “has the potential to direct learners’ attentional and memory resources to the way the L2 structures and codes concepts” (Robinson, 2011a, p. 15), thus negatively affecting fluency, but positively influencing accuracy and complexity. On the other hand, cognitive task complexity along resource-dispersing dimensions can be manipulated in various ways. For example, the provision of pre-task planning can decrease cognitive task complexity along resource-dispersing dimensions. Removal of planning time, however, can make a task more performatively complex, dispersing learners’ attention over many linguistic and non-linguistic aspects of the task. Increasing cognitive complexity of tasks along resource-dispersing dimensions “promotes not noticing of language code, and interlanguage development of new linguistic, conceptual form-function mappings”, but “consolidation and fast real-time access to existing interlanguage resources” (Robinson, 2011a, p. 17).

Task condition includes participation and participant variables. Participation variables mean “nature of the solution to, and direction of information flow on, tasks” (Robinson, 2007a, p. 14), for example, whether the solution to the tasks is open or closed, and whether information exchanges are one-way or two-way. Participant variables refer to “how different configurations of participants affect the amount and quality of interaction” (Robinson, 2007a, p. 14), such as gender, and familiarity.

Task difficulty refers to learners’ perceptions of the cognitive demands made by the task; it is not relative to task features, but the abilities and affective responses that learners bring to the task. Robinson (2007a) highlighted the importance of learner factors in task complexity, as it “is essential to optimising the chances of success for learners, and for program level decisions about individualising instruction” (p. 19). According to Robinson (2007a, 2011a), affective variables, such as anxiety, a predictor of perceived task difficulty, will increasingly differentiate learning and performance as tasks increase in complexity, since affective factors may “change and so affect the size of resource pool availability on a temporary basis” (Robinson, 2001b, p. 32).

Across task complexity, task conditions, and task difficulty, Robinson proposed that task complexity should be the sole basis for sequencing decisions, as it can be controlled largely by teachers or syllabus designers regardless of learner or situation factors. Task difficulty and task conditions, which cannot be predicted before a course starts, have little to contribute in terms of task sequencing, but they can inform the online-decisions about how task should be implemented in the classroom.

Robinson (2010) suggested that tasks should be sequenced from simple to complex, based on task complexity dimensions and the SSARC model to foster L2 development. In the SSARC model, “SS” stands for “simple/stabilizing interlanguage”. “A” means “automatizing access to interlanguage”. “RC” refers to “restructuring and complexifying” (Robinson, 2010, p. 252). Robinson proposed that the first task should be simple along the resource-directing and resource-dispersing dimensions, which is called the SS stage. Then, at the A stage, tasks should be made more cognitively complex along the resource-dispersing dimension, such as removing pre-task planning time. Tasks can be made even more cognitively demanding by involving little prior knowledge that learners have. In this way, the performative demands to access existing linguistic resources are gradually increased, promoting real-time access to the interlanguage resources. Finally, at the RC stage, the complexity of tasks should be increased along the resource-directing dimension to further stretch, restructure, and complexify learners’ current level of interlanguage system.

Like Skehan (2009), Robinson (2011b) draws on Levelt’s L1 speaking model to provide the theoretical rationale for how task demands affect L2 speech performance. Robinson (2011b) argued that greater effort at the conceptualization stage, induced by increases in the conceptual demands of tasks, can promote the “development and re-mapping of conceptual and linguistic categories” (p. 16) at the lexcio-grammatical encoding stage, with positive consequences for accuracy and complexity in learners’ performance. Although they both draw on Levelt’s (1989) model, Skehan and Robinson diverge in their interpretation and the hypotheses that they derive from the model.

In CH, Robinson (2011a) makes five ancillary theoretical claims on the potential effects of task complexity on language learning and production. Firstly, increasing cognitive task complexity along resource-directing dimensions in monologic tasks will lead to simultaneous enhancement in accuracy and complexity, but decrease in fluency. Secondly, cognitively complex tasks will contribute to more interaction and negotiation of meaning, and promote learners to learn from the input and incorporate forms that are made salient in the input. Thirdly, cognitively demanding tasks will promote “greater depth of processing" and then "longer-term retention of input" than the simple tasks. Fourthly, tasks sequencing from simple to complex will lead to "greater automaticity and efficient scheduling of the components of complex L2 task performance" (Robinson, 2011a, p. 19) than sequencing in other ways, such as from complex to simple. Finally, learners' affective and ability variables will be more influential on learners' perceptions of task difficulty, and more clearly differentiate learners' performance in complex tasks than simple tasks. Moreover, Robinson has hypothesized that simultaneously increasing task complexity along resource-directing dimensions, and reducing the complexity of resource-dispersing task characteristics, will affect learners' L2 performance positively.

In summary, both Skehan's LACM and Robinson's CH focus on the effects of task complexity on oral language production and predict how attentional resources are allocated during task completion. The two models make the same prediction for the effects of task complexity along resource-dispersing dimensions (in Robinson's terminology); that is, increasing the cognitive demands of tasks along resource-dispersing dimensions would negatively affect learners' oral language performance in terms of CAF.

The two models, however, differ in two areas: (a) the amount of attentional resources, and (b) predictions of the effect of increasing task complexity along resource-directing dimensions on language output. The LACM holds that learners have limited attentional resources and can allocate attention to only one dimension of complexity, accuracy, or fluency at a time when task complexity is increased along resource-directing dimensions. By contrast, Robinson assumes that learners have multiple attentional resources and that dimensions of task complexity belong to different and non-competitive attentional resource pools. Increasing task complexity along resource-directing dimensions, therefore, will simultaneously enhance complexity and accuracy, but reduce fluency.

Skehan's LACM and Robinson's CH are two influential oral theories, which have been frequently used in L2 writing task-based research. The present study, situated within these two cognitive task complexity models, explores one of the most important questions in L2 writing pedagogy: How do different task characteristics influence the quality of students' writing?

2.2 Kellogg's Writing Model

Kellogg's (1996) writing model is famous for the initial prediction of the cognitive load placed by writing processes in relation to working memory. The model is depicted in Figure 2.1. Kellogg (1996) distinguished three different components of the writing process with each comprising two sub-processes: formulation (*planning* [1] and *translating*), execution (*programming* and *executing*), and *monitoring* (*reading* and *editing*).

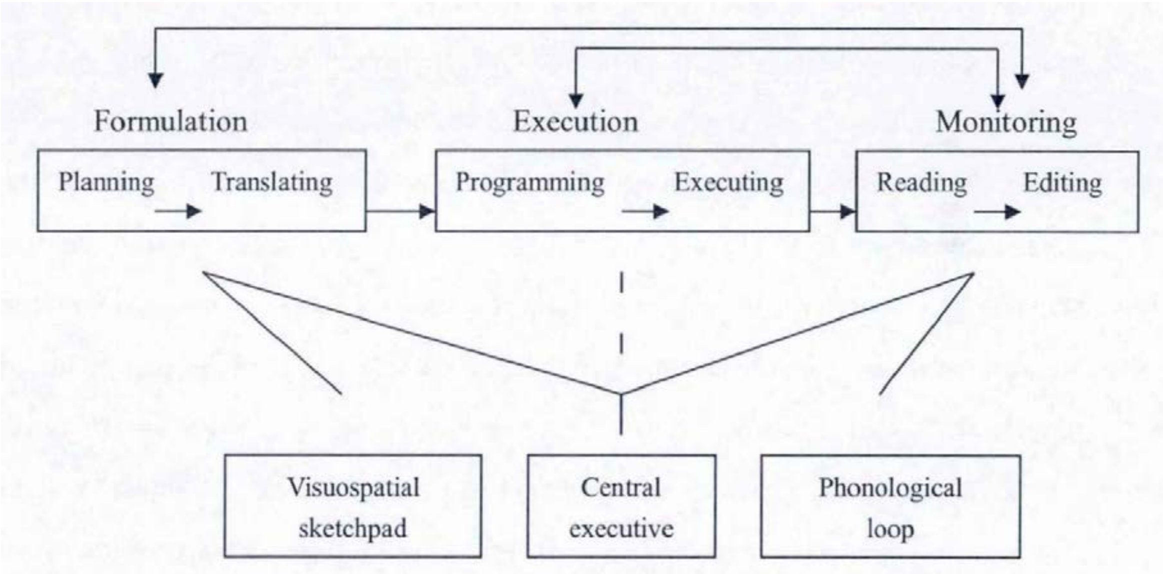

Figure 2.1 Kellogg's model of writing processes (reprinted from Ellis & Yuan, 2004, p. 62)

The formulation system plans the content and translates ideas into sentences; during *planning*, writers use their working memory to establish goals and retrieve ideas and knowledge to write. Translation is "the amalgam of linguistic processes needed to convert an idea into a written message" (Kellogg, 1996, p. 60); the lexical, syntactic and rhetorical items are selected to encode ideas into words during *translating*. In the execution system, while *programming* refers to adopting the appropriate motor system (e.g., handwriting, typing or dictating) to translate the output, *executing* occurs when the production is created into words and sentences based on the translating process. *Monitoring* refers to *reading* and *editing* the produced version to ensure the writer's

[1] The processes, such as *planning, translating, and monitoring* in Kellogg's model are all italicized to distinguish students' action, such as do planning and translating.

intention is adequately expressed. During the *editing*, writers can either edit the localized errors, like spelling and diction, or the global problems, such as the paragraph and text organisation issues.

Figure 2.1 shows the flow of information in the model between the basic processes and the entire system. Writing information, as indicated by arrows, flows from formulation to execution and forward to *monitoring*. For example, the input of translation can be the output of *planning*, and the output of *translating* may then be sent to the execution system. The model allows for the simultaneous activation of formulation, execution, and *monitoring* and their interaction, as long as the demands do not exceed the limitations of the central executive capacity. The bidirectional arrows indicate that *monitoring* can occur both before and after the execution process. For example, outputs of *planning* and *translating*, such as ideas and writing goals, can be fed forward to *editing* before a sentence is programmed and executed. A new round of formulation might be prompted when the previous ideas or goals planned are rejected by *editing*. Corrections may be made after handwriting, typing, or dictating, if execution takes place before *editing*.

According to Baddeley (1986), there are two slave systems of the central executive in the working memory: The phonological loop that stores and processes the auditory and verbal information, and the visuo-spatial sketchpad that includes the storage and processing of the visual and spatial information. The two slave systems can operate independently, but if the task demands become too great for slave system, the central executive can be activated to aid successful performance.

Kellogg (1996) proposed, based on Baddeley's (1986) constructs, the central executive is a crucial component of working memory, which is involved in the production of texts. The central executive is involved in all the sub-processes except executing. The formulation system places the heaviest demands on both the central executive and the slave components. "Planning in all of its facets engages the central executive" (Kellogg, 1996, p. 63). During *planning*, writers engage the visuo-spatial sketchpad when they visualize ideas, organize schemes, and invoke their visual imagery when "creating ideas and recalling them from long-term memory" (Kellogg, 1996, p. 62). Translating an idea into a written message involves the phonological loop, and *translating* requires the central executive resources "when the writer must struggle to find just the right words and sentence structures" (Kellogg, 1996, p. 63). The execution system makes demands only on the central executive, since typing and handwriting presumably place light demands when the skills are well practiced. The dotted line in Figure 2.1 presents the minimal demands of well-practiced execution. With the *monitoring* system, *editing* “takes so many forms, ranging from the detection of a motor programming error to a revision in the organization of ideas in a text” (Kellogg, 1996, p. 65), and makes heavy demands on the central executive, but not on the phonological loop. *Reading* requires the resources of both the phonological loop and the central executive, although the demand on the central executive by *reading* is less than *editing*. To summarize, in the three-process writing production model, the formulation system is theorized to place the greatest demands on working memory capacity, whereas execution, which is theorized to be relatively automatic, places minimal demands on working memory.

Kellogg et al (2013) evaluated the 1996 Model based on a large number of current studies on working memory and writing. They confirmed that “*planning*, *translating*, and *reviewing* are all dependent on the central executive” (Kellogg et al., 2013, p. 170), and that *translating* requires demands on verbal working memory and the central executive. Additionally, the assumption was extended that the planning of sentences for concrete words, rather than abstract language, evokes visual imagery. Finally, the assumption that *editing* requires no demands on the phonological loop was seriously challenged.

Kellogg et al. (2013) suggested three instructional ways to reduce overload on the working memory of developing writers during composition, based on the 1996 model. Firstly, having expertise in a domain-specific topic allows writers to reduce the temporary demands made on executive attention and to “retrieve relevant knowledge from long-term memory at just the right moment” (Kellogg et al., 2013, p. 184). Kellogg (2013) claimed that writing a familiar topic allows writers to devote more executive attention to juggling *planning*, *generating* and *monitoring*, instead of overloading working memory, than the case when writing on an unfamiliar topic. This position is consistent with the perspective of “familiarity of topic/task” in Skehan’s LACM (Skehan, 1998; Skehan & Foster, 1999) and “prior knowledge” in Robinson’s (2001b, 2005, 2011a, 2011b) CH. Secondly, attentional resources can be funnelled to a single process at a given moment in time, benefiting the quality and fluency of writing, by using some strategies. For example, the use of outlining during pre-task planning helps writers funnel more attention to *planning* than *monitoring*, and to the macrostructure than microstructure. As a result, writers can free their working memory resources to translate ideas into linguistic codes during writing. By contrast, writers may share their limited central executive capacity to manage the coordination of *planning*, *translating* and *monitoring* under time constraints. This assertion is similar to both Skehan's and Robinson's claim that the provision of pre-task planning contributes learners' better performance. Finally, "extensive and deliberate practice" (Kellogg et al., 2013, p. 185) can prepare writers to achieve relatively automatic processing in *planning*, *translating* and *monitoring*, consequently lessening the burden of each writing process on working memory. Deliberate practice requires: "(1) effortful exertion to improve performance, (2) intrinsic motivation to engage in the task, (3) practice tasks that are within reach of the individual's current level of ability, (4) feedback that provides knowledge of results, and (5) high levels of repetition" (Kellogg et al., 2013, p. 185).

Kellogg's model neither considers the role of task environment and affective factors in writing, nor predicts performance outcomes of manipulating task design and implementation factors. The model does, however, provide insight into writers' processes in relation to working memory. Although Kellogg's model is designed for L1 writers, it can also apply to L2 writers, who may experience more pressure when writing in L2 than L1 writers. With limited central executive capacity, L2 writers have to compromise among three writing processes when under pressure to produce text rapidly. It can be hypothesized that varying planning types and task design features affect L2 writers' attentional allocation to different writing processes during writing. In the present study, Kellogg's (1996) model will be used to explain the effects of two types of planning, pre-task planning and online planning, relative to working memory allocation.

2.3 Cognitive Task Complexity in Relation to Writing Process

Writing, a recursive process, is different from speaking, a linear process (Kellogg, 1990). Kellogg's (1996) model provides a foundation for theorizing the influences of cognitive task complexity on L2 learners' written performance.

2.3.1 Cognitive-task variability in relation to writing process

Cognitive-task variability of EFL writing tasks “can be hypothesized to be inherent in the formulation stage and can be determined by the demand tasks make on the planning of the content of the written text and/ or on the linguistic encoding of the content” (Kormos, 2011, p. 150). To elaborate, increasing the conceptual demands of tasks, which are inherent in the task, leads to greater effort to conceptualize the content at *planning* of writing process. Consequently, complex linguistic structures are used to encode complex concepts, augmenting the demands of *translating* during writing. In this way, more cognitively complex tasks may demand more cognitive attentional resources in both *planning* and *translating* stages of writing. Furthermore, Kormos (2011) argued that it is possible that tasks make separate complexity demands on the *planning* and *translating* processes, since varying manipulations of the task features might have diverse effects on the writing process. For example, Kormos (2011) hypothesized that a task with a fixed story line may ease students’ cognitive demands on conceptualization in *planning*, but exert pressure during *translating* to encode the prescribed content. Whereas, students, when completing the task without a story line, may use more cognitive load on the content conceptualization in *planning* at the cost of allocating less attention to *translating,* as they can tailor the content to match their linguistic resources.

2.3.2 Planning type in relation to writing process

Kellogg’s model provides a basis for identifying the writing components L2 learners focus on during planning and writing. As Ellis (2005) concluded, rehearsal, a type of pre-task planning, may assist writers to attend in both formulation and *monitoring*, leading to all-round improvements. Strategic planning, a type of pre-task planning, may help learners mainly focus on conceptualisation, which contributes to greater message complexity and increased fluency. Unpressured online planning may enable learners to devote attention to formulation and *monitoring*, resulting in a more accurate performance. Ellis and Yuan’s (2004) findings confirmed the prediction that pre-task planning and online planning impact different aspects of EFL writing processes, with the former promoting formulation and the latter providing better opportunities for *monitoring*.

In this study, therefore, the cognitive task complexity hypotheses and Kellogg's (1996) writing model are used as the basis from which to explain how varying cognitive-task variability manipulations and different planning types affect EFL learners' writing process in relation to working memory.

2.4 Empirical Studies into Cognitive Task Complexity in EFL Writing along Resource-directing Dimensions

Cognitive-task variability, in the present study, refers specifically to the resource-directing dimensions of the cognitive task complexity, concerning the variables making cognitive/conceptual demands on participants' attention (Robinson, 2005, 2007a). Cognitive-task variability in previous studies has been operationalized in various ways to explore the role of cognitive task complexity in EFL writing field. The operationalisation of cognitive-task variability in the field of EFL writing, along with some representative studies for each are summarized in Table 2.4. Previous studies, where the cognitive-task variability was operationalized as [±Here-and-Now], [±reasoning demands], [±casual reasoning] and [±content support], are reviewed in Section 2.4.1. In Section 2.4.2, studies related to the present one, in which cognitive-task variability was manipulated by [±few elements], are introduced.

Table 2.4 Summary of Cognitive-task Variability Operationalisation with Example Studies in EFL Writing

Cognitive-task variability operationalization	**Example studies**
±Here-and-Now	Ishikawa (2007)
±Reasoning demands	Ruiz-Funes (2014, 2015)
±Casual reasoning	Yang, Lu & Weigle (2015)
±Content support	Kormos (2011), Révész, Kourtali & Mazgutova (2017), Abrams (2019)
± **Few elements**[2]	Kuiken, Mos, & Vedder (2005), Kuiken & Vedder (2007b, 2008, 2012), Frear and Bitchener (2015), Cho (2015), Rahimi (2019), Rahimi & Zhang (2019), Kang and Lee (2019)

2.4.1 Effects of cognitive-task variability: [±factors][3]

The absence versus presence of content has been one of the most frequently researched resource-directing variables of cognitive task complexity in the field of L2 writing. Kormos (2011) designed two narrative tasks and varied the cognitive load by manipulating the plot. For a simple cartoon description task, a comic strip consisting of six pictures with a coherent story line, was provided. In contrast, the complex task was a narrative task based on six unrelated pictures. The secondary-school participants had to conceptualize the story-line in their imagination and create the story with their language skills. No significant effects of cognitive-task variability on linguistic performance were found, except for a major effect on lexical sophistication and a minor effect on temporal cohesion, which are in contrast to the predictions of CH.

Later, Révész et al. (2017) explored the effects of cognitive-task variability, manipulated by [±content support], on 73 advanced L2 participants' writing behaviours and linguistic complexity. Less pausing, more revision and greater linguistic complexity (i.e.,

[2] This column was highlighted in bold, as the cognitive-task variability in the present study was manipulated by ±few elements

[3] Factors include: ±Here-and-Now, ±Reasoning demands, ±Casual reasoning, ±Content support, listed in Table 2.4.

increased use of less frequent words and greater overall syntactic complexity) were identified when content support was provided, whereas the complex task, without content support, led to more frequent pauses and revisions and less sophisticated lexis. The findings for lexical complexity are consistent with those of Kormos (2011), although the task type was different; narrative writing tasks were used in Kormos' (2011) study, while argumentative writing tasks were used in Révész et al.'s (2017) research. Révész et al. (2017) speculated that content support in this study may have helped reduce participants' processing burden, allowing them to allocate more attentional resources to linguistic encoding.

Recently, Abrams (2019) launched a long-term study to explore the effects of cognitive task complexity on L2 writing performance. Twenty-three intermediate learners of German were asked to complete six writing tasks over ten weeks, among which three were content supported integrated writing (simple tasks) and three were non-supported writing (complex tasks). This study found that the simple tasks yielded significantly better writing performance in terms of complexity, grammatical and lexical accuracy, fluency and lexical richness. The results lend support to Kormos' (2011) and Révész et al.'s (2017) findings and confirm the prediction that the availability of content support may reduce learners' cognitive processing burden.

Ishikawa (2007) manipulated cognitive-task variability in terms of [±Here-and-Now]. With the cognitively simple task [+Here-and-Now], 27 secondary-school participants, who viewed a cartoon strip during writing were asked to write in the present tense, while their 27 counterparts with the cognitively complex task [-Here-and-Now] were required to memorize and retrieve the cartoon-strip and write in the past tense. Learners in the complex task produced greater syntactic complexity, accuracy and fluency in their writing, which supports for the CH, but contradicts the findings of Kormos' (2011), Révész et al.'s (2017), and Abrams' (2019) studies. The difference in the results, however, may be because the participants were given 5 minutes of preparation time before writing (Skehan, 2009).

In Ruiz-Funes' (2015) study, cognitive task complexity was manipulated by reasoning demands and familiarity of topic, genre, and task type. College-level FL writers of Spanish at two proficiency levels were recruited as participants to explore the effects of task complexity on linguistic performance as well as a mediating factor of language proficiency. For advanced learners, the simple task was an analytical writing task, while the complex one was an argumentative writing task. For intermediate learners, the simple task asked students to write a personal essay, while the complex task required an expository essay. The simple tasks were intended to be cognitively less demanding, as they required students to rely on their own personal experiences and a familiar genre, while the complex ones were assumed to be more cognitively demanding in terms of topic, genre, and task type. At each proficiency level, participants were divided into sub-levels of performance based on their overall writing quality. The results showed that within the given levels of language proficiency, increased cognitive task complexity led to an increase in syntactic complexity, but a slight decrease in accuracy and fluency. A tendency towards trade-off among measures of linguistic production lends support to the principles in Skehan's LACM. Ruiz-Funes (2015) also argued that trade-off effects tend to be diminished when participants are at more advanced language proficiency levels and higher performance levels in writing, which is consistent with Robinsons' CH.

Yang, Lu and Weigle (2015) investigated the effect of different topics on syntactic complexity in L2 writing within the framework of cognitive task complexity. Two argumentative writing topics were used: One, the appearance topic "whether people place too much emphasis on personal appearance", has fewer causal reasoning demands, while the other is the future topic, "whether careful planning while young ensures a good future" (Yang et al., 2015, p. 57), which requires more causal reasoning. One hundred and ninety graduate students' 380 essays were analysed from Weigle's (2011) database. As for the dependent variables, syntactic complexity was assessed in a multi-dimensional way with eight different measures used to represent the eight inter-connected sub-constructs: overall sentence complexity (MLS), clausal coordination (T/S), overall T-unit complexity (MLT), clausal subordination (DC/T), elaboration at clause level (MLC), phrasal coordination (CP/C), noun phrase complexity (CN/C) and non-finite elements/subordination (NFE/C). The results showed that the future topic, as the complex task with causal reasoning demands, yielded a higher amount of subordination as measured by DC/T and NFE/C and greater global sentence complexity as measured by MLS. Whereas, the appearance topic, as the simple task without causal reasoning demands, elicited more elaboration at the finite clause level as measured by MLC, CP/C and CN/C. The results confirm Ruiz-Funes' (2015) prediction that "the certain topics may naturally call for more use of certain local-level syntactic complexity features" (p. 63).

Taken together, findings of previous studies (Ishikawa, 2007; Kormos, 2011; Ruiz-Funes, 2015; Yang et al., 2015) on the effects of cognitive-task variability on EFL writing are inconclusive. The divergence of results, however, may be due to the different operationalisations of cognitive-task variability.

2.4.2 Effects of cognitive-task variability: [±few elements] factor

The most commonly operationalized cognitive-task variability in L2 writing research, as shown in Table 2.4, is to change the number of elements involved in the tasks. According to Robinson (2011a), increasing the number of elements entailed in the tasks can augment the cognitive-task variability, simultaneously facilitating accuracy and complexity, but negatively affecting fluency. Whereas, Skehan (1996a) hypothesized that increases in cognitive-task variability with more elements limit learners' attention to one aspect of complexity, accuracy, or fluency.

Kuiken and his colleagues (Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012) conducted a series of studies on the effects of cognitive-task variability[4] on L2 narrative letter-writing tasks. The topic in the three studies was about choosing a holiday destination from options. In the complex task, a choice of a Bed and Breakfast from five options had to be made based on six criteria (six elements), while in the simple task writers had to choose a holiday resort in a distant country based on three requirements (three elements). Participants were required to complete the tasks within 40 minutes with a minimum of 150 words. The use of a dictionary was allowed. Linguistic performance was assessed by means of syntactic complexity (C/T, DC/C)[5], lexical variation (WT/W, WT/√2W, and/or lexical sophistication, or TTR), and accuracy (Err/T, and/or specific types of errors).

[4] Other independent factors, such as proficiency and mode, and relative findings in these studies were not reviewed in this part.

[5] C/T=clauses per T-unit; DC/C=dependent clauses per clause; WT/W=the number of word types divided by the total number of word tokens; WT/√2W = the total number of different word types divided by the square root of two times the total number of words; TTR=Type-token Ratio; Err/T=the number of errors per T-unit.

In Kuiken et al's (2005) study, 62 Dutch university students, learning Italian as L2 in their first-, second- and third-year, were recruited. The results showed that increased cognitive-task variability yielded a significant positive effect on accuracy but no significant effects on syntactic and lexical complexity. In Kuiken and Vedder's (2007b) study, the participants were 84 Dutch university students, who were beginners in learning Italian, and 75, who had studied French as L2 for a couple of years. The results showed both students of Italian and French had greater accuracy in terms of lexical errors in the complex task than the simple task, but there were no significant differences in syntacitic complexity between the two task versions. As for lexical performance, findings for students of Italian and French differed. Students of Italian used signficantly more high frequent words, while the students of French used more infrequent words in the complex task than in the simple task. In Kuiken and Vedder's study (2008), 91 Dutch university students of Italian (beginners) and 76 learners of French (with 5- or 6-years experience) were recuited as participants. For both language learning groups, accuracy in the complex task was signficantly greater than in the simple task, while no signficant differences in syntactic complexity and lexical variation were identified between the simple and complex task versions. In Kuiken and Vedder's (2012) research, three studies were conducted to investigate the influence of cognitive-task variability on linguistic performance. In Study 1, increased cognitive-task variability did not significantly influence participants' syntactic complexity and lexical variation, but positively affected accuracy. In Study 2, in addition to the general measures used in Study 1, specific measures for lexical frequency, and accuracy in terms of Grammar, Lexicon, Spelling and Appropriateness were also employed. Results showed that the complex task was performed more accurately by the students of both Italian and French, due to fewer lexical errors produced, than the simple task. The students of Italian and French showed different patterns of lexical variation, with the students of Italian using significantly more high frequent words and students of French more infrequent words in the complex task than in the simple task. In Study 3, increasing cognitive-task variability led to more accuracy, but had no effect on lexical variety and syntactic complexity.

To sum up, studies conducted by Kuiken and his colleagues (Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012) consistently revealed that increased cognitive-task variability had a positive effect on accuracy, which partially support Robinson's CH.

No significant effects on syntactic complexity were, however, identified as a function of increasing cognitive-task variability. The researchers explained that results may indicate that learners primarily allocate their attentional resources to controlling their existing interlanguage system rather than complexifying linguistic encoding. The findings of the lexical performance in these studies, however, were mixed, although they shared the same cognitive-task variability manipulation and similar proficiency level of participants.

Later, Frear and Bitchener's (2015) research with 34 non-native speakers of English in an Auckland school partially replicated studies of Kuiken and Vedder (2007b, 2008, 2012). They introduced an apparently low complexity task as the independent variable to better track the effects of cognitive-task variability. Cognitive-task variability was manipulated at three levels. Task 1, the simple task, asked participants to write a letter to a friend to introduce New Zealand, based on their own resources, which was expected to have a limited cognitive demand. Tasks 2 and 3, with medium and high level task complexity respectively, required participants to choose a restaurant for visiting friend(s) based on the information about each restaurant and the friend's requirements. Task 3 was more cognitively complex than Task 2 because more elements were entailed in the tasks, one more option for the restaurants, and two more visiting friends with their preferences. As for the dependent variables, the ratio of dependent clauses was analysed as a whole and separately (i.e., adjectival, nominal and adverbial dependent clauses) to measure more precisely the syntactic complexity. Additionally, a mean segmental type-token ratio (MTTR) was used to measure lexical complexity. The results indicate that lexical complexity improved as a function of increasing cognitive-task variability, especially between Task 1 and 3. No significant differences in syntactic complexity were identified, when the ratio of dependent clauses was analysed as a whole. Nonetheless, significant decreases in ratio of adverbial dependent clauses to T-units, were found as cognitive-task variability increased: Task 1 had a higher adverbial dependent clauses/T-units than Tasks 2 and 3. These findings for syntactic complexity differed from previous studies (Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008), although they shared a similar operationalization of independent and dependent variables. In summary, the favourable effects on lexical complexity appears to lend support to Robinson's CH. With the decrease, or no change, in syntactic complexity, there is an apparent trade-off between lexical and syntactic means of expression, suggesting that participants may prioritise their limited cognitive resources to “the easier lexical means of meeting the pragmatic requirements of the tasks” over “the generation of grammar” (Frear & Bitchener, 2015, p. 53).

To date, the most frequently used task type is narrative writing, with the number of studies on argumentative writing relatively scarce. Compared to the narrative genre, argumentative writing requires more critical thinking and articulation of ideas, and more challenging use of linguistic elements (Hirose, 2003), thus requiring greater cognitive effort of writers.

Cho (2015) explored the effects of cognitive-task variability on argumentative writing, a departure from narrative tasks (Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012). Cho (2015) manipulated the cognitive-task variability by varying the number of argument elements and reasoning demands within the topic of “choosing the best roommates”. In the simple task, participants were asked to choose one pair of roommates from four options, each marked by four properties. In the complex task, participants had to select two pairs from six candidates, each marked by six properties. The linguistic measures included both global (i.e., EFT/T and Err/T for accuracy, C/T and T/S for syntactic complexity, and W/T and W/S for fluency) and task-specific measures (i.e., frequency of conjunctions for syntactic complexity, and target-like use of conjunctions for accuracy). The results showed that increased cognitive-task variability significantly positively affected fluency, but did not significantly influence accuracy and syntactic complexity in either global or specific measures. The findings contradict the prediction of CH, but support Skehan’s LACM. Cho (2015) argued that participants’ prioritization of fluency, according to Skehan (1996a), may interfere with attention to the dimensions of accuracy and complexity. The results of Cho’s (2015) study differed from previous findings with narrative tasks (Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012), due possibly to the different task types (i.e., narrative versus argumentative writing); the influence of task types on learners’ attention may contribute to the variable findings (Kuiken et al., 2019; Skehan, 2009).

Recently, Rahimi and Zhang (Rahimi, 2019; Rahimi & Zhang, 2019) studied the effects of cognitive-task variability by manipulating the number of elements and reasoning demands on argumentative writing production in an Iran EFL learning context. With a topic of project funding allocation, the simple task required participants to allocate \$5,000, 000 across three competing projects, while in the complex task, a \$10,000,000 fund was to be allocated across six competing projects. Quality of the text (content, organisation and overall score) was measured. Traditional aspects of linguistic performance were also assessed: syntactic complexity (i.e., MLC, subordination and phrasal coordination); lexical complexity in terms of lexical diversity (i.e., values of D) and lexical sophistication (i.e., academic words); and accuracy (i.e., EFT/T, Err/T, EP100 and Weighted Clause Ratio). The results demonstrated that increased complexity of the task led to enhancements in syntactic complexity, lexical complexity, L2 content, organization and writing quality, which partially supports the predictions of CH. A significant negative impact on accuracy was, however, found. The trade-off between complexity and accuracy lends support to Skehan's LACM.

To conclude, mixed findings were found in previous studies (Cho, 2015; Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007a, 2008, 2012; Rahimi, 2019; Rahimi & Zhang, 2019), although cognitive-task variability in these studies was identically manipulated by [±few elements]. The divergent results may be in large part due to the variety in task types and linguistic measures employed as well as learner factors (Kormos, 2011; Kuiken & Vedder, 2008; Norris & Ortega, 2009; Ong & Zhang, 2010; Rahimi & Zhang, 2019).

2.5 Empirical Studies into Cognitive Task Complexity in EFL Writing along Resource-dispersing Dimensions

2.5.1 Effects of planning type

Skehan (Skehan, 1996a, 1998, 2009, 2015, 2018; Skehan & Foster, 1999) and Robinson (Baralt et al., 2014b; Robinson, 2001b, 2005, 2011a, 2011b; Robinson & Gilabert, 2019) made the same prediction for the effects of cognitive task complexity with resource-dispersing dimensions; that is, decreasing the cognitive demands of tasks along resource-dispersing dimensions facilitates learners' CAF performance. Among the resource-dispersing variables, planning time, especially pre-task planning, has been more extensively researched than planning type in studies of L2 speaking (Ellis, 2003, 2009; Foster & Skehan, 1999; Ortega, 1999; Skehan, 2018) and writing (Abrams & Byrd, 2016; Ellis & Yuan, 2004; Johnson, Mercado, & Acevedo, 2012; Ong & Zhang, 2010, 2013). In studies on written language, more researchers have focused on pre-task planning (Abrams & Byrd, 2016; Johnson et al., 2012; Ong & Zhang, 2010, 2013) than online planning (Ellis & Yuan, 2004). Previous studies on pre-task planning and on online planning in the field of L2 writing are reviewed in this section.

Ong and Zhang (2010) recruited 108 Chinese university students to explore the effects of task complexity on L2 learners' argumentative writing production. Task complexity was manipulated by three independent variables: varying planning conditions, $\pm$ ideas and/or macrostructure support and $\pm$ draft available during revision[6]. Four planning conditions were examined with the total time controlled: Extended pre-task planning (20 minutes for planning and 10 minutes for writing), pre-task planning (10 minutes for planning and 20 minutes for writing), free-writing (write immediately and continuously for 30 minutes), and control group (write in their own way for 30 minutes). Based on Robinson's theory, researchers supposed that task complexity increased from extended pre-task planning, to pre-task planning and then to free-writing conditions. Three conditions were designed based on ideas and macrostructures: topic, ideas and macro-structure given, topic and ideas given, and topic given. Participants' writing performance was analysed in term of lexical complexity (WT^2/W)[7] and fluency (fluency I calculated for writing time and fluency II for total time). The results showed that increased task complexity with pre-task planning time decreased from 20 to 0 minute(s) led to significantly greater fluency II and lexical complexity. Increased task complexity with no support of ideas or macro-structure resulted in significantly greater lexical complexity, but there was no significant influence on either fluency I or fluency II. This study, however, did not report the relationship among the traditional three dimensions of linguistic performance (i.e., CAF), with fluency and lexical complexity but no accuracy or syntactic complexity involved.

Ong and Zhang (2013) used the same independent variables as their earlier study (Ong & Zhang, 2010). Ong and Zhang (2013) explored how planning, idea related sub-planning and revising conditions influenced L2 learners' writing quality, which was rated using Jacobs, Zinkgraf, Wormuth, Hartfiel and Hughey's (1981) criteria. The findings showed that free-writing conditions unexpectedly facilitated participants'

[6] Revision conditions, unrelated to planning, are not reviewed in this part.
[7] Word types squared divided by the total number of words.

writing quality. The authors posited that free writing may assist learners' content retrieval, which enhanced the text quality. Moreover, the increased task complexity with topic given resulted in decreased text quality, compared to the conditions of idea given, as well as idea and macro-structure given.

In sum, the two studies (Ong & Zhang, 2010, 2013) found that pre-task planning may inhibit writers' fluency, lexical complexity and writing quality. This result was contrary to either findings of previous studies in the field of L1 writing (Kellogg, 1987, 1990), or prior findings of EFL oral production (Skehan & Foster, 1997). The small sample size in each group in the two studies may be a reason for the mixed results.

Motivated by the contradictory findings, Johnson, et al. (2012) investigated the effect of different sub-processes of pre-task planning on fluency, grammatical complexity and lexical complexity in argumentative writing. Nine hundred and sixty-eight EFL learners with Spanish as their first language were recruited as participants. Participants were given 10 minutes for pre-task planning and 30 minutes for writing. The pre-task planning was operationalized as five sub-conditions: Idea generation planning (to list as many ideas on the topic as possible), organisation planning (to outline main ideas, supporting ideas or examples), goal setting planning (to list rhetorical goals), goal setting and organisation planning (to list goals and make an outline) and control group (to complete a vocabulary matching activity during the planning time). The results showed that pre-task planning had some inconsequential effects on fluency, with participants in organisation planning condition producing shorter sentences than those in the control group. Pre-task planning, however, had no impact on lexical complexity or grammatical complexity. These findings are in contrast not only the positive findings that pre-task planning improved L2 learners' writing fluency and complexity (Ellis & Yuan, 2004), but also the results that pre-task planning negatively impacted L2 writing fluency, lexical complexity and writing quality (Ong & Zhang, 2010, 2013). Johnson, et al. (2012) speculated that LACM and CH, which were proposed for oral tasks, may not be accurately appropriate for writing scope. The distinction between speaking and writing may have led to inconsistent findings in the studies on pre-task planning in L2 writing field. Individual learner factors and language proficiency also need to be taken into consideration.

In Abrams and Byrd's (2016) study, 26 first-year university learners of German were recruited to perform three written summary tasks (two with and one without pre-task planning) to explore the effects of pre-task planning on L2 writing performance. With the two tasks preceded by planning, participants completed either a mind-mapping or chronological sequencing planning, which were counterbalanced. The dependent variables included syntactic complexity, grammatical accuracy, fluency, lexical choice, lexical accuracy and propositional content. Pre-task planning, regardless of planning type, had a significant favourable effect on lexical complexity, fluency and propositional content. Learners who completed the task without planning had greater lexical accuracy, more accurate lexical choices and error-free t-units than those with any types of pre-task planning. The positive effect of pre-task planning on fluency is consistent with Ellis and Yuan (2004), but inconsistent with the results found in other prior studies (Johnson et al., 2012; Ong & Zhang, 2010). Of note is that the result that no-planning condition benefited accuracy is in contrast with Ellis and Yuan's (2004) findings. These results lend support to Skehan's LACM, as "participants' performance demonstrated either better accuracy (grammatical and lexical) or richer content (fluency, propositional content, lexical richness), but not both" (Abrams & Byrd, 2016, p.8).

In contrast to these quantitative studies, Ojima (2006) used a case-study approach to investigate the impact of concept mapping as a form of pre-task planning on three Japanese ESL learners' writing performance. The three participants were asked to write four essays, two with pre-task planning and two without. During the pre-task planning phase, 10 minutes was allocated for concept mapping. Written fluency, complexity, and accuracy were measured. The four writing tasks were not operationalized within the cognitive task complexity framework. The results showed a positive effect for pre-task concept mapping on fluency and complexity, but not on accuracy. From the qualitative data, the differences in the application of concept mapping on participants' writing processes were identified. Ojima (2006) argued that pre-task concept mapping may help ESL learners improve their writing performance in ways unique to individual learners.

Ellis and Yuan (2004) compared the effects of pre-task planning, unpressured online planning and no planning on EFL narrative written performance. In the pre-task planning condition, participants were given 10 minutes to plan and 17 minutes to write. In the online planning condition, no time limit was set, which meant that participants could take as long as they liked to complete the task. By contrast, only 17 minutes were provided for participants in the no planning group. Fluency in terms of syllables per minute and number of dysfluencies, complexity as measured by syntactic complexity, syntactic variety and MSTTR, and accuracy in terms of error-free clauses and correct verb forms were measured as the dependent variables. The results showed that pre-task planning led to greater fluency, while unpressured online planning resulted in greater accuracy. In contrast, the no-planning condition negatively influenced fluency, complexity and accuracy of written production. The results suggest that planning conditions affected participants' writing process differently. Pre-task planning facilitated formulation process, while unpressured online planning assisted *monitoring* process. In contrast, writers with no-planning time had to deal with formulation, execution, and *monitoring* under pressure. The results, however, showed pre-task planners (27 minutes in total) spent more time on the task than online planners (21 minutes on average) and no-planning writers (17 minutes). The uncontrolled total time may act as confounding variables and consequently affect the task performance. This study, as the previous one was not framed within the cognitive task complexity.

In summary, most prior studies paid attention to pre-task planning, while the few studies with a focus on online planning were not conducted under the umbrella of cognitive task complexity. The present study investigates the effects of pre-task planning and online planning on EFL argumentative writing within the framework of cognitive task complexity and examines whether the two planning types differently affect learners' writing process.

2.6 Empirical Studies into Cognitive Task Complexity in EFL Writing Simultaneously along Two Dimensions

Few studies, to date, have systematically examined the combined effects of task complexity along resource-directing and resource-dispersing dimensions (Gilabert, 2007; Rahimi & Zhang, 2018), although "in real-world communicative tasks these two dimensions simultaneously affect performance" (Kormos, 2011, p. 151).

Gilabert (2007) explored the synergistic effects of task complexity along these two dimensions on oral production. Cognitive-task variability was operationalized in terms of [±Here-and-Now], and only pre-task planning was adopted as a resource-dispersing factor. The findings showed that the provision of pre-task planning generated a higher level of fluency, and lexical richness, as well as higher, but non-significant, structural complexity and accuracy than the no-planning condition for both simple and complex tasks. Whereas, increasing task complexity along [±Here-and-Now] reduced fluency, lexical complexity and structural complexity, but increased accuracy under both planned and unplanned conditions. Gilabert (2007) concluded that simultaneous attention to complexity and accuracy may be achievable with negative consequences for fluency, if tasks are simple along resource-dispersing (i.e., providing planning time) and complex along resource-directing (i.e., -Here-and-Now) dimensions. These results confirm Robinson's predictions that "if tasks are kept simple along resource-dispersing variables but are made more complexity along resource-directing variables, attention may be allotted to complexity and accuracy simultaneously" (cited in Gilabert, 2007, p. 63). This research provides direction for exploring the synergistic effects of cognitive-task variability and planning on EFL performance, although it focuses on EFL oral production.

Recently, Rahimi and Zhang (2018) extended Gilabert's (2007) study by examining the synergistic effects of task complexity along two dimensions in the field of EFL writing. Eighty Iranian learners of English at an upper-intermediate level, recruited as participants, completed two argumentative writing tasks. Task complexity along resource-directing dimensions was manipulated by varying numbers of elements and reasoning demands. The topic of the two writing tasks was about the allocation of two sources of funds. Resource-dispersing dimensions, with or without the provision of 10-minute pre-task planning time were manipulated to decrease or increase cognitive task complexity. Participants' writing performance was assessed for syntactic complexity, lexical complexity, accuracy, fluency, content, organisation, and overall writing quality. The results showed that providing pre-task planning time produced more fluency, syntactic complexity, content, and overall writing quality, but without significant effects on accuracy or lexical complexity in either simple or complex tasks. Increases in cognitive-task variability generated greater syntactic complexity, lexical complexity, content, and overall writing quality, but reduced learners' accuracy and fluency under planning and no planning conditions. Rahimi and Zhang (2018) concluded that simultaneously increasing cognitive-task variability and providing pre-task planning improved participants' syntactic complexity in terms of subordination, content, organisation, and overall writing quality. Joint enhancement in complexity and accuracy was not achieved, since increasing cognitive-task variability impeded learners' accuracy and providing pre-task planning had no effects on improvement in accuracy. These results partially support Robinson's predictions that simultaneously increasing task complexity along resource-directing dimensions and reducing the complexity of resource-dispersing task characteristics may positively affect learners' L2 performance.

Taking Gilabert's (2007) and Rahimi and Zhang's (2018) studies together, a general pattern of improvement in learners' performance was apparent when task complexity was manipulated by increasing resource-directing variables and decreasing resource-dispersing variables. The relationships between complexity and accuracy were, however, still inconsistent, possibly due to the small number of studies that differ in research context, task modality, operationalisation of cognitive-task variability, or varying production measures.

Recently, Kang and Lee (2019) explored the effects of pre-task planning types (i.e., individual planning and collaborative planning) on L2 narrative writing, particularly in relation to task complexity. In contrast to most previous studies that manipulated [±planning time] as a resource-dispersing factor of task complexity, pre-task planning type in this study was treated as a within-subjects variable. Task complexity, as a between-subjects variable, was manipulated in terms of [± few elements] with one main character in the simple tasks and more than three characters in the complex tasks. Forty 8th-grade Korean EFL learners at were assigned to the simple the complex task groups, with 20 participants in each group. Participants in the simple task group performed two simple tasks: The first with individual planning and second with collaborative planning. When the first task was finished, participants could select their partners and discuss their story with their partner during the collaborative planning, but they had to finish the stories individually after the collaboration. Six minutes were provided for both planning types but with no detailed instructions. During the two types of planning, participants could write notes, but the notes were removed before writing. The procedures for the complex task group were the same, except that participants completed two complex tasks and were given seven minutes planning time. No time limit was set for completing the story in both groups. Lexical complexity (i.e., MTLD, and the log frequency of content words), syntactic complexity (i.e., MLT, and MLC), accuracy (i.e., E/T) and fluency were the dependent variables assessed. Kang and Lee (2019) reported that using collaborative planning led to greater fluency and syntactic complexity than individual planning, regardless of task complexity. In the complex task group, the two types of pre-task planning resulted in a similar level of lexical complexity, while for the simple task group, individual planning provided an apparent advantage over collaborative planning for lexical complexity. The findings suggest that collaboration at the pre-task stage can facilitate L2 learners' writing in syntactic complexity at global and phrasal levels, and that "individual or collaborative planning should be determined in light of task complexity" (Kang & Lee, 2019, p. 70). The non-significant effect of task complexity on students' writing performance in Kang and Lee's (2019) study is inconsistent with either previous empirical studies on task complexity (Cho, 2015; Kuiken & Vedder, 2012; Rahimi, 2019; Rahimi & Zhang, 2018, 2019; Ruiz-Funes, 2015), or the predictions of CH (Robinson, 2001b, 2005, 2011a, 2011b) and LACM (Skehan, 1996a, 1998, 2003, 2018). The difference in the research outcomes may be attributable to participants' time allocation during planning and writing, as well as participants' different levels of English proficiency and implementations of pre-task planning. The complex task group was given an additional minute, which may favour participants' performance, and no time limit for completing the tasks may have allowed for online planning during writing. Consequently, the extent to which online planning was involved in students' writing process may obscure the effects of pre-task planning and task complexity. Online planning, again, was not included as resource-dispersing variable in this study.

Based on the previous studies reviewed above, the findings on the effects of cognitive-task variability on EFL writing are far from conclusive even where these studies shared the same manipulation of cognitive-task variability as [±few elements] (Cho, 2015; Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008; Rahimi, 2019; Rahimi & Zhang, 2019). Few studies have investigated the effects of both pre-task planning and online planning on L2 writing from the perspective of cognitive task complexity, and in the small number of studies on the simultaneous effects of task complexity along the two dimensions on EFL writing, the results are mixed.

Through careful review of the previous studies, possible reasons for the mixed results for cognitive task complexity can be concluded as the following. Firstly, **the cognitive-**

task variability was not validated prior to the study. Instead, the researchers assumed that the complex task consumed more cognitive resources than the simple task (Baralt et al., 2014a; Révész et al., 2016; Sasayama, 2016). Secondly, **various but not multi-dimensional linguistic measures, were employed** in the studies, which may have led to different findings (Housen et al., 2019; Kuiken et al., 2019; Norris & Ortega, 2009; Pallotti, 2009; Rahimi, 2019; Yang et al., 2015). Thirdly, **individual differences,** as highlighted in Robinson's CH, **may impact the way that cognitive task complexity influenced EFL writing performance**, a variable which has been called for by many researchers (Ellis & Yuan, 2004; Kormos, 2011; Kuiken & Vedder, 2008; Ong & Zhang, 2010; Rahimi & Zhang, 2019). These three feasible reasons are discussed in detail in the following sections.

2.7 Validation of Cognitive-task Variability

Independently measuring every construct, according to Norris and Ortega (2003), is vital for evaluating any theoretical framework. Within the paradigm of cognitive task complexity, researchers have traditionally assumed, rather than proved, that the designed task versions distinguished levels of cognitive demands. For example, cognitive-task variability, a key construct of cognitive task complexity, has been typically operationalized by designing simple and complex versions. The complex version has usually been assumed to cost more cognitive load than the simple one. To date, most studies on cognitive task complexity, however, have not assessed whether different levels of cognitive demands were, in fact, achieved between task versions (Révész et al., 2016).

Cognitive demand, or cognitive load, as a multidimensional construct, reflects "the interaction among task-generated processing demands and learner abilities" (Révész et al., 2016, p.707). The assessment of cognitive load requires multiple measurements (Révész, 2014; Révész et al., 2016; Révész, Sachs, & Hama, 2014; Sasayama, 2016).

Techniques for measuring cognitive task demands have been actively used in the area of cognitive psychology, although little attention has been paid to the issue in instructed SLA research (Révész, 2014). According to Brünken, Seufert and Paas (2010), current techniques can be divided into two categories: (a) subjective approaches, such as, self-ratings and time estimations, and (b) objective approaches, such as, dual-task methodology and some behavioural physiological measures (e.g. heart rate and eye movement). Of these techniques, self-ratings and dual-task methodology are reviewed in the following sections.

2.7.1 Self-rating scales as measures of cognitive load

Self-rating scales may be the most commonly used method to measure cognitive demands in the field of educational psychology (Brünken et al., 2010) and in SLA (Révész et al., 2016). With this measurement, learners are required to assess their cognitive load by answering a semantic differential scale, such as “I invested…mental effort on this task”, with options varying from “very, very low” to “very, very high” (Brünken et al., 2010, p. 182). The rationale for self-rating scales is that learners can reliably and validly evaluate the amount of cognitive load they spent on a specific cognitive activity (Brünken et al., 2010). Additionally, researchers usually combined the cognitive load self-rating scale with some indirectly related subjective ratings of variables, such as task difficulty and motivation, to form a multidimensional assessment tool. These variables, such as self-perceived task difficulty, however, do not have identical constructs with perceived cognitive load or mental effort, although they are often highly correlated (Brünken et al., 2010). For example, cognitive load or mental effort refers to “the amount of resources actually allocated to accommodate the task demands” and “the amount of controlled processing in which the individual is engaged” (Paas, van Merriënboer, & Adam, 1994, p. 420). Task difficulty refers to learners’ perceptions of the cognitive demands made by the task, which is determined by the abilities and affective responses that learners bring to the task (Robinson, 2005, 2007a, 2011a). Mental effort focuses on the actual effort made by individuals, while perceived task difficulty emphasizes the task itself. The multidimensional scale, with a combination of mental effort and task difficulty, can provide a comprehensive, and scientific, evaluation of perceived cognitive load.

Robinson (2001b) is acknowledged as one of the earliest studies in the field of task-based research to independently measure cognitive-task variability by using a self-perceived task difficulty scale. This self-rating questionnaire assessed participants’ perceived levels of task difficulty, stress, confidence, interest, and motivation using a 9-point Likert scale. The results showed that participants rated a designed-to-be simple task as significantly less difficult and stressful than the designed-to-be complex task.

Although this self-assessment scale has been widely and actively used in the recent task-complexity studies (Rahimi, 2019; Rahimi & Zhang, 2018, 2019; Robinson, 2007b), these studies do not include items that tap self-perceived mental effort.

To date, only Sasayama (2016) and Révész et al. (2016) have used a scale that includes mental effort and task difficulty in the field of task-based research. A 2-item, paper-based questionnaire was first used by Sasayama (2016) to measure both task difficulty and mental effort: "(a) Was telling the story in English easy? (b) How much brain power did you use to complete this tasks?" (p. 238). The options provided ranged from "very easy/not at all" to "very difficult/intensively" on a 9-point Likert scale. She found that among the four oral tasks, the simplest task was rated as the least difficult and the least mental-effort-demanding, while the most complex one rated as the most difficult and requiring the greatest mental effort. These results are consistent with those yielded by the other approaches used in the research, dual-task methodology and time estimation. The results of self-ratings of task difficulty and mental effort on the four tasks may not, however, accurately represent the cognitive load yielded by the four oral tasks per se, as participants completed each task under a dual-task condition. The results of self-ratings on task difficulty and mental effort, therefore, may include not only participants' perceptions of performing the oral task, but also their perceptions related to the dual-task condition.

Révész et al. (2016) used a 6-item computer-based questionnaire on a 9-point Likert scale to measure the manipulations of the simple and complex versions of three pairs of oral tasks: A pair of picture narrative tasks, a pair of map tasks, and a pair of decision-making tasks. In contrast to Sasayama's (2016) research design, participants in Révész et al.'s (2016) study were allocated into two groups, a dual-task condition and a single condition (without an added secondary task). Only the students in the single condition completed the subjective scales after each oral task. The findings showed that for the map and decision-making tasks, but not the narrative tasks, the designed-to-be complex version was rated as more difficult and consumed more mental effort than the designed simple task version. The results for self-ratings were found to be aligned with their intended task manipulations.

To conclude, the prior two studies (Révész et al., 2016; Sasayama, 2016) indicated that the multidimensional self-assessment is a viable way to assess the task-generated cognitive demands. More research in the field of EFL writing, however, is warranted to test the utility of self-rating scales in confirming the manipulations of cognitive-task variability. As suggested by previous studies (Brünken et al., 2010; Révész et al., 2016; Sasayama, 2016), additional objective measures are needed to supplement subjective self-rating measures.

2.7.2 Dual-task methodology as a measure of cognitive load

Dual-task methodology, as suggested by its name, involves performing an additional secondary task simultaneously with a primary task. During the primary task, participants react to a variable-timed probe stimulus (i.e., secondary task) as quickly as possible. Dual-task methodology, as an objective measurement of cognitive load, serves as a direct measure of cognitive load without influence of mediating variables, and is independent from the learners' self-report (Brünken, Plass, & Leutner, 2003; Schoor, Bannert, & Brünken, 2012).

The principle underlying the dual-task technique is that the amount of cognitive load induced by the primary task will be reflected in the secondary task performance, which is normally assessed in terms of reaction time and accuracy rate (Brünken et al., 2003; Révész, 2014; Révész et al., 2016, 2014; Schoor et al., 2012). For instance, when participants complete a primary task with more cognitive demands, less accuracy and more reaction time are expected on their secondary task performance than when they perform a simple primary task. According to the suggestions of several researchers (e.g., Brünken et al., 2003; Marcus, Cooper, & Sweller, 1996; Schoor et al., 2012), secondary tasks should be in the form of simple activities, such as, an easy visual or auditory stimulus; provide as little disturbance as possible to the primary task; be continuously scored, independent of the primary task performance; and not suppress the quality or quantity of primary task performance.

Dual-task methodology has been employed in a great number of cognitive psychology studies (Brünken et al., 2003, 2010; Marcus et al., 1996; Schoor et al., 2012), and is considered a reliable and sensitive approach for measuring cognitive load, but has received little consideration in the field of SLA (Révész, 2014).

Recently, a few researchers have started to adopt dual-task methodology in studies of cognitive task complexity. For example, Révész and her colleagues conducted a series of studies using dual-task methodology, and other techniques, to independently assess the validity of cognitive-task variability. In Révész, Sachs, and Hama's (2014) study, the two oral primary tasks had different reasoning demands. Twenty-four participants (16 native speakers of English and 16 EFL university students) described a famous person's life by using a past counterfactual statement based on two past events they chose. The secondary task was to respond to the change of the computer background screen, which was programmed to randomly change the color to red or green for 250 milliseconds during each 2,500-millisecond interval. Half of the participants were instructed to respond to green (ignoring red), while the other half were asked to respond to red (ignoring green). Accuracy and reaction time were recorded to reflect the level of cognitive load imposed by the primary tasks. The findings showed that participants produced more accuracy rates on the secondary task when performing the simple version of the primary task, although no significant differences were found in the secondary-task reaction time. Later, Révész, Michel and Gilabert (2016) triangulated results from dual-task methodology with other techniques (i.e., self-rating scale and expert judgment) to respectively measure the cognitive load imposed by the primary tasks. The secondary tasks used in these two studies (Révész et al., 2016, 2014) were identical, although the primary tasks were quite different. Three pairs of simple and complex oral tasks were used as the primary tasks and the three modes were picture narrative, map, and decision-making tasks. Participants achieved a higher mean accuracy rate but non-salient reaction time when carrying out the simpler version of the primary tasks. Similar patterns for accuracy and reaction time on the secondary task were found in both studies. Révész et al. (2016) postulated that the accuracy rates might be more sensitive than reaction time under the colour detection secondary-task condition.

Sasayama (2016) employed the dual-task method, together with time estimation and self-rating techniques, to measure cognitive-task variability. She selected a colour-detection design as the secondary task. The colour of letter A was programmed to change unpredictably from black to red and black to black. The red colour appeared during each 5,000-milliseconds interval, while the duration of the black letter was set with a variation from 3,000 to 60,000 ms. Participants were required to react as quickly as possible only when letter A was changed from red to black. Four narrative tasks, which involved an increasing number of elements, were designed to vary the cognitive demands of the primary tasks. Sasayama found a trend that participants responded to the colour stimulus the fastest when carrying out the designed-to-be-simplest task, and the slowest when engaging in the most complex one, although this difference was not significant. The same pattern, however, was not identified for the intermediate complex versions. Sasayama (2016) reported consistent patterns of results on three measurements of cognitive-task variability.

Révész et al. (2016, 2014) and Sasayama (2016) pioneered the dual-task paradigm to independently measure the cognitive load of the oral tasks in task-based research. The primary tasks in these three studies are oral, and the secondary tasks are presented in a visual way. Different secondary tasks are linked to different steps of information processing, reflecting the different processes of the cognitive load imposed (Baddeley, 1986; Brünken et al., 2003). It is suggested that the secondary tasks are chosen based on the modality of the primary tasks (Schoor et al., 2012). The visual secondary task with a colour detection design is obviously not appropriate when the primary tasks are written. Reed, Burton, and Kelly's (1985) study is one of the very few that have used dual-task methodology in the field of language writing research. Their study explored the effects of writing ability, and mode of discourse, on cognitive capacity engagement by employing dual-task methodology. Sixty-three college freshmen of varying writing abilities (i.e., basic, average, and honours) were instructed to perform the dual-task methodology. Three writing tasks in forms of descriptive writing, narrative writing and persuasive writing served as the primary tasks. The secondary task was to react to a tone at random from a microcomputer by pressing the space bar. The tone was presented 20 times in total, randomly appearing five times at the prewriting stage, 10 times at the writing stage and five times at the rewriting stage. Cognitive-capacity engagement was analysed by participants' reaction time. The results showed that writers' cognitive-capacity engagement across different modes was differentially affected by their writing ability. Reed et al.'s (1985) study, to some extent, provides a different viewpoint on using dual-task methodology in the field of writing research, although it is not relevant to the cognitive task complexity.

In sum, the dual-task technique has been used infrequently to measure the cognitive load imposed by the writing tasks in the field of task-based research. The auditory stimulus can be applied in the secondary task for the dual-task paradigm, as in Reed et al.'s (1985) research, to measure the cognitive load imposed by primary writing tasks.

2.8 Measurements of EFL Writing Performance

The most commonly used measures to assess learners' writing performance in task-based research in the field of SLA are the linguistic criteria in terms of syntactic complexity, lexical complexity, accuracy, and fluency. These linguistic measures can be divided into task-specific and global measures.

Task-specific measures can assist to examine the sub-constructs of each dimension of CAF, which are "inter language-sensitive" (Robinson, 2005, pp. 21-22). Some previous studies have used such measures to capture the effects of task complexity on learners' writing performance (Cho, 2015; Kuiken & Vedder, 2007b, 2012). Task-specific measures, however, may be influenced by the topics or goals, instead of the manipulation of task complexity (Pallotti, 2009), which will not provide accurate information in general nor be generalizable to task-based research.

Global linguistic measures used by most previous task-based studies (Cho, 2015; Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008; Rahimi, 2019; Rahimi & Zhang, 2018, 2019) to assess learners' writing performance will be reviewed. Fluency measures, which have not been controversial, will not be discussed in this section.

2.8.1 Syntactic complexity

Syntactic complexity refers to "the range of forms that surface in language production and the degree of sophistication of such forms" (Ortega, 2003, p. 492). As researchers suggested (Housen et al., 2019; Kuiken et al., 2019; Norris & Ortega, 2009), syntactic complexity should be measured multi-dimensionally to tap distinct and complementary dimensions of complexity. At least three measureable sub-constructs were suggested to measure syntactic complexity: (a) overall complexity, (b) complexity via subordination to capture the intermediate proficiency level complexity, and (c) phrasal elaboration to capture the syntactic complexity development of an advanced proficiency. According to Halliday and Mattiesen (1999), learners develop their ability to use phrasal-level complexification later than hypotaxis (i.e., subordination). Subordination has been, therefore, considered a useful index of complexification at an intermediate level. Phrasal-level elaboration, as a pervasive means to measure participants' syntactic complexity, is an index of an advanced level in the field of writing (Biber, Gray, &

Poonpon, 2011; Housen et al., 2019; Norris & Ortega, 2009). In addition to these measures, complexity via coordination should also be included, if early proficiency data are involved (Norris & Ortega, 2009).

Most previous studies, as reviewed in Sections 2.4, 2.5, and 2.6, used only two or three syntactic measures, which sometimes are redundant of the same sub-constructs of syntactic complexity. For example, the three commonly used measures, which were found correlated to EFL learners' proficiency, are: the number of clauses per T-unit (C/T), the number of dependent clauses per T-unit (DC/T) and the number of dependent clauses per total number of clauses (DC/C) (Wolfe-Quintero, Inagaki, & Kim, 1998). These three measures, however, only capture the linguistic development in subordination dimensions of syntactic complexity (Norris & Ortega, 2009).

Only a small number of previous studies (Rahimi, 2019; Rahimi & Zhang, 2018, 2019; Yang et al., 2015; Yoon, 2017) used a multi-dimensional syntactic measures to explore the effects of task complexity. Among them, Rahimi and his colleague (Rahimi, 2019; Rahimi & Zhang, 2018, 2019) assessed three sub-constructs without including an overall measure of syntactic complexity. Yoon's (2017) study was not framed within task complexity. It may be the case that some certain dimensions of syntactic complexity may be improved with corresponding increases in task complexity, but the lack of comprehensive measures employed in previous studies may not have captured such enhancements. Multi-dimensional measures of syntactic complexity are, therefore, suggested in the future to explore the effects of cognitive task complexity on EFL linguistic performance (Housen et al., 2019; Kuiken et al., 2019; Norris & Ortega, 2009; Pallotti, 2009; Rahimi, 2019; Yang et al., 2015).

2.8.2 Lexical complexity

Lexical complexity is "manifested in writing primarily in terms of the range and size of a second language writer's productive vocabulary" (Wolfe-Quintero et al., 1998, p. 102). As a multidimensional feature of a learner's language use, it has been measured widely in terms of lexical density, lexical diversity, and lexical sophistication (Lu, 2012). Little prior research, to date, has assessed all three dimensions in the field of task-based research. Among the three dimensions, lexical diversity is the commonly used index, which refers to "the range of a learner's vocabulary as displayed in his or her language use” (Lu, 2012, p.192). Type-token Ratio (TTR) has been the most popular method used to address lexical diversity. “Type” means the number of different words and “token” means the total number of words. The calculation result of TTR is in a proportion from 0 to 1, with a higher figure showing a more diverse range of vocabulary in a given sample. TTR has been criticised for the lack of sensitivity to text length. To overcome this shortcoming, TTRs that took text length into account, were created, like the index of Guiraud (WT^2/W) or its variant: $WT/\sqrt{2W}$ (Carroll, 1967), which were widely used in previous studies (Ishikawa, 2007; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012; Ong & Zhang, 2010).

Lexical density refers to “the ratio of the number of lexical (as opposed to grammatical) words to the total number of words in a text” (Lu, 2012, p. 191). According to Lu (2012), lexical words were defined as “nouns, adjectives, verbs (excluding modal verbs, auxiliary verbs, ‘be’ and ‘have’), and adverbs with an adjective base” (p. 192). This index of lexical complexity has rarely been used in prior research to explore the effects of task complexity on lexical complexity performance.

Lexical sophistication, a measure for “the proportion of relatively unusual or advanced words in the learner’s text” (Read, 2000, p. 203), has recently been used in a number of previous studies (Kormos, 2011; Kuiken & Vedder, 2007b; Rahimi, 2019). The use of the lexical frequency profile, however, varied across studies (Kormos, 2011; Kuiken & Vedder, 2007b; Rahimi, 2019). For instance, Kormos (2011) measured participants’ vocabulary range with the help of three bases: “Nation’s Range program”, “Ageneral Service List of English Words” and “The Academic Word List”, and found more abstract words were elicited by a complex task. In Kuiken and Vedder’s (2007b) research, the ratio of sophisticated word types was calculated based on Laufer and Nation’s (1995) lexical frequency profile. In the complex task, students of Italian used more high frequent words and students of French used more infrequent words. Finally, Rahimi (2019) found participants used significantly more academic vocabulary on the complex task than the simple task, by measuring the use of academic vocabulary and less frequent vocabulary based on Coxhead’s (2002) profile. Without a unified lexical frequency profile for the lexical sophistication, different topics and task requirements may easily result in varying outcomes of lexical sophistication.

To conclude, a multi-dimensional and consistent picture of learners' linguistic complexity performance influenced by task complexity is not yet available, since previous studies commonly assessed a few measures of syntactic and/or lexical complexity at one or two dimensions in the task-based research in the field of EFL writing. An extensive amount of calculating work would be needed for incorporation of a multi-dimensional measures of syntactic and lexical complexity. To reduce the amount of work, some computer-calculated analysers of syntactic complexity and lexical complexity have been developed, such as Coh-Metrix (McNamara, Louwerse, McCarthy, & Graesser, 2010), EFL Syntactic Complexity Analyzer (EFL SCA; Lu, 2010) and Lexical Complexity Analyzer (EFL LCA; Lu, 2012). The EFL SCA (Lu, 2010) with 14 measures of syntactic complexity and LCA (Lu, 2012) with 25 kinds of measures have been proven with a high validity and accuracy and has been increasingly used in the studies on L2 writing in the task-based research (Rahimi, 2019; Rahimi & Zhang, 2018, 2019; Yang et al., 2015).

2.8.3 Accuracy

Accuracy refers to "how well the target language is produced in relation to the rule system of the target language" (Skehan, 1996b, p. 23). Skehan and Foster (1999) suggested that differences between experimental conditions can be more sensitively detected by a general measure of accuracy. Wolfe-Quintero, Inagaki, and Kim (1998) recommended three measures for accuracy: the number of error-free T-units (EFT), error-free T-units per T-unit (EFT/T) and the number of errors per T-unit (Err/T). The first two measures were useful for advanced learners (Kuiken & Vedder, 2007b). EFC/C was a popular measure for L2 beginners and (low) intermediate learners, since it is not always easy to find any error-free T-units in their L2 writing performance. Apart from Err/T, EP100 was also used to assess the number and type of errors.

2.8.4 Adequacy

Most studies have assessed CAF within writing tasks in the field of task-based research with very few (Ong & Zhang, 2013; Rahimi, 2019; Rahimi & Zhang, 2018, 2019) focusing on the issue of how a topic unfolded and whether task goals have been successfully achieved. Learners' performance may be extremely high on complexity, accuracy, and fluency, but irrelevant to the pragmatic requirements of a task (Pallotti, 2009). A separate dimension, adequacy, which is "theoretically independent from CAF"

(Pallotti, 2009, p. 596), therefore, needs to be tested for the pragmatic appropriateness in the field of task-based writing research.

Adequacy represents “the degree to which a learner’s performance is more or less successful in achieving the task’s goals efficiently” (Pallotti, 2009, p. 596). In studies reviewed in Sections 2.4, 2.5 and 2.6, only Ong and Zhang (2013) assessed overall text quality and Rahimi and Zhang (2018, 2019) examined content, organisation, and overall writing quality, with all based on Jacobs et al.’s (1981) rating criteria. Additionally, explicit predictions concerning the effects of cognitive task complexity along either dimension on adequacy could not be made from CH and LACM frameworks. The effects of cognitive task complexity on adequacy is needed to extend the task-based research both empirically and theoretically.

2.9 Anxiety in Effects of Cognitive Task Complexity

As a complex, multidimensional phenomenon, language anxiety has been addressed as a separate term from a general construct of anxiety (Horwitz, Horwitz, & Cope, 1986). Second language anxiety is defined by MacIntyre (1999) as “the apprehension experienced when a situation requires the use of a second language with which the individual is not fully proficient…the propensity for an individual to react in a nervous manner when speaking, listening, reading, or writing in the second language” (p. 5) (see also recent studies, e.g., MacIntyre, 2017, MacIntyre, Gregersen, & Mercer, 2016). Language anxiety is considered as one of the most important affective factors, obstructing the success of language learning (Arnold, 2000; Horwitz, 2001, 2016; MacIntyre, 2017; Sheen, 2008; Zhang, 2000, 2001), which is also highlighted as one of the individual differences factors by Robinson (2011a) in CH. A substantial body of studies that have explored the relationship between language anxiety and various broad-based L2 achievement, such as course grades (Horwitz & Young, 1991; MacIntyre & Gardner, 1994), have found three patterns of language anxiety effects: a debilitating effect, facilitating effect and no effect (Dörnyei, 2005). Few studies, however, have explored the relationship between language anxiety and the process of learning and production in L2 acquisition, especially in the field of L2 writing.

Tobias’ (1986) model for studying the cognitive effects of anxiety on learning, was pioneering in this area. According to Tobias (1986), language anxiety cognitively affects

L2 pedagogic tasks at three stages: Input, Processing and Output stages. Input is the initial stage where instructions are presented to students. In this early stage, highly anxious students are more likely to divide their attention between task demands and irrelevant stimuli (e.g., worries), which could inhibit all subsequent stages, with fewer items stored in memory for later retrieval. Processing occurs after external instructions have been encoded and subjected to cognitive processing. During this stage, students engage in planning the message and making linguistic preparations to meet the task goals. According to Tobias (1986), cognitive processing will be impeded by anxiety at the Processing stage if the tasks are difficult with heavy demands on memory or are poorly organized. Output refers to the completion of tasks processed at previous stages. “It is at this stage that language learners are required to demonstrate their ability to use the second language” (MacIntyre & Gardner, 1994, p. 287). The impact of language anxiety on performance at the Output stage has been widely explored. The three stages are not isolated, but rather interdependent, and sometimes occur simultaneously in the real-life learning cases. “Language anxiety affects each stage of the cognitive processing of language production in such a way that anxiety experienced at one stage causes increased levels of anxiety at subsequent stages” (Trebits, 2016, p. 157).

Inspired by Tobias’ (1986) model, MacIntyre and Gardner (1994) investigated the effects of language anxiety on specific cognitive processes in the second language. A set of nine tasks, representing the Input, Processing and Output stages, were employed by MacIntyre and Gardner (1994) to isolate and measure different stages of language acquisition. They found that anxiety disrupts the cognitive processes by diverting students’ attentional resources away from the task at each of the three stages and that these effects seem cumulative. For example, at the Input stage, a heightened state of anxiety may interfere with the language learning, which will further reduce anxious students’ language knowledge base at the subsequent stage. Finally, students’ retrieval for appropriate linguistic items during the Output stage may be inhibited and slowed down, especially under a time constraint. MacIntyre and Gardner’s (1994) result is in line with those found in the cognitive psychological field that anxiety negatively affects cognitive performance, especially on tasks that require attentional focus (Derakshan & Eysenck, 2009; Eysenck, Derakshan, Santos, & Calvo, 2007).

Using MacIntyre and Gardner's (1994) Input, Processing and Output (IPOA) scale, Robinson (2007b) explored the role of anxiety plays in the effect of cognitive task complexity on EFL oral production. Forty-two Japanese L1 university students were employed as speaker/storyteller or listener/sequencer to complete three interactive oral tasks with cognitive-task variability manipulated by reasoning demands. Participants' anxiety at Input, Processing and Output stages were assessed at the very beginning of the study using IPOA scale. The study found that participants' oral production of the task with heavier cognitive demands was more significantly affected by Output anxiety than the simpler version of the task, thus confirming the CH that "individual differences in ability and affective factors relevant to the cognitive demands of tasks will increasingly differentiate learners' speech production…as tasks increase in complexity" (Robinson, 2007b, p. 196).

Later, Trebits (2016) compared spoken and written output relative to the effect of anxiety at the Input, Processing and Output stages, and cognitive task complexity. Forty-four second-year students with a bilingual educational background were recruited in this study. Two types of narrative tasks (i.e., carton description with clear cues and picture narration without a coherent storyline) in speech and in writing were manipulated as simple and complex tasks. The IPOA scale was used to measure students' level of anxiety. The result showed that the effects of Output anxiety on participants' spoken output are significantly stronger than those on written modality. Input and Processing anxiety, unexpectedly, affected lexical and structural complexity positively but negatively influenced participants' accuracy performance, which was identified as a trade-off between accuracy and complexity. Trebits (2016) speculated that highly anxious students may find it more difficult to balance their attention between complexity and accuracy, since "anxiety may impair the regulation of attentional resources and attentional control during performing cognitive operations" (p. 170).

Révész (2011) used two argumentative tasks on the topic of allocating funds for public projects (i.e., a simple version and a complex version) to explore the effect of anxiety, self-confidence and self-perceived communicative competence on cognitive-task variability on L2 oral performance. Forty-three L2 university participants were recruited. A three-item anxiety questionnaire adapted from Kormos and Dörnyei (2004) was used to assess participants' anxiety level. The results demonstrated that anxiety had no significant association with learners' oral performance, which not only rejects Robinson's (2011a) CH, but also contradicts the findings obtained by MacIntyre and Gardner (1994), Robinson (2007b) and Trebits (2016). Révész (2011) argued that the absence effect of anxiety may be due to the advanced proficiency level of participants who were probably able to use strategies to minimize the adverse impact of anxiety.

Previous studies (Révész, 2011; Robinson, 2007b; Trebits, 2016) have focused mainly on oral language production, while limited research has investigated the intertwined effects of cognitive task complexity and language anxiety on written production. Although Trebits' (2016) study involved written output, the IPOA scale was not a suitable assessment of students' writing anxiety. Only 1 of the 18 items in the scale explicitly applied to writing anxiety. The remaining 17 items addressed anxiety in terms of speaking, listening, readings and testing. This is one possible reason for the weak effects of anxiety on writing performance found in Trebits' (2016) study.

Writing anxiety, as a situation- and subject-specific language anxiety, has been found to be negatively correlated with writing performance (Cheng, 2004; Cheng, Horwitz, & Schallert, 1999; Daly & Miller, 1975a, 1975b; Horwitz, 2001). Writing in a second language is a difficult task with great challenges for L2 learners which easily arouses writers' anxiety (Cheng, 2004; Woodrow, 2011; Zhang, 2013). To independently study and assess L2 writing anxiety by using a specific, and multi-dimensional, writing anxiety scale is necessary, especially for the issue of how writing anxiety interferes with cognitive task complexity in the writing process.

Rahimi and Zhang (2019) is leading the field in relating writing anxiety to the effects of task complexity on EFL writing. They used Cheng's (2004) *Second Language Writing Anxiety Inventory* (SLWAI) to assess participants' L2 writing anxiety (other details have been reviewed in section 2.4.2). The SLWAI, which has been validated in different contexts (Choi, 2013; Guo & Qin, 2010; Rahimi & Zhang, 2019), demonstrates high validity and reliability. L2 writing anxiety is multidimensional and consists of three components in the SLWAI: Somatic Anxiety, Avoidance Behaviour and Cognitive Anxiety. Rahimi and Zhang (2019) found that in the complex task, only Avoidance Behaviour had a moderate negative relationship with the mean length of T-unit and with academic vocabulary use, while no relationship between anxiety and writing performance was shown in the simple task, which is consistent with Robinson's (2011a) CH.

The small set of studies on the relationship of anxiety with cognitive task complexity on L2 students' (speaking or writing) performance were correlational. Students' anxiety level has usually been assessed by their completion of a questionnaire. A qualitative approach, such as interview, has seldom been used "to better understand the role of anxiety in language learning" and learners' "feelings about language learning" (Yan & Horwitz, 2008, p. 153). The role anxiety played in the effects of cognitive task complexity on L2 performance is still far from consistent among the limited number of studies (MacIntyre & Gardner, 1994; Révész, 2011; Robinson, 2007b; Trebits, 2016). Additionally, little attention, to date, has been given to the effects of writing anxiety on cognitive task complexity in the process of writing. The relationship between resource-dispersing characteristics of task complexity (i.e., planning) and writing anxiety has also been under-researched.

The present study seeks to extend the existing research in the L2 writing field by investigating the relationship among L2 writing anxiety, cognitive task complexity along both resource-directing and resource-dispersing dimensions, and L2 argumentative writing performance. In addition to the L2 writing anxiety questionnaire, a semi-structured interview was designed to elaborate the data collected from the questionnaire and, more importantly, to track participants' writing anxiety when they engage in varying conditions at different writing stages.

2.10 Chapter Summary

In this chapter, Skehan's LACM and Robinson's CH were introduced. Kellogg's (1996) Writing Model and Kellogg et al.'s (2013) Evaluation of the 1996 Model were described. These existing models, however, have not addressed the issue of how cognitive load, imposed by task complexity, affects EFL writing process and performance.

The literature review highlights the following gaps on which there is little research: Firstly, the previous findings on the effects of cognitive task complexity along either resource-directing (i.e., ±few elements) or resource-dispersing (i.e., planning type) dimensions on EFL writing are inconclusive, which indicates the need for sustained investigations in each dimension of cognitive task complexity. Furthermore, little attention has been paid to the simultaneous effects of cognitive task complexity. The inclusion of different planning types, especially online planning, various topics of writing tasks, an enlarged sample size, and different EFL learning contexts, is warranted to extend the understanding of the simultaneous effects of cognitive task complexity along the two dimensions on EFL writing.

Secondly, although the effects of cognitive-task variability on EFL writing have received increased attention, little research in the field of EFL writing has provided independent evidence for validating the cognitive-task variability, a pivotal construct in cognitive task complexity. Unlike self-rating scales, the dual-task technique, an actively used cognitive psychology method, has rarely been used to measure the cognitive demands imposed by writing tasks in the field of task-based research. The present study, therefore, is an initial effort to use multiple independent measures of cognitive load for triangulation (i.e., dual-task method and self-rating) to validate the manipulations of cognitive-task variability in the field of EFL writing.

Thirdly, most previous studies used redundant but not multi-dimensional sub-constructs of CAF to examine the effects of cognitive task complexity on EFL writing. Moreover, the issue of how a topic unfolded and whether pragmatic goals have been successfully achieved have hardly been addressed. Multi-dimensional measures of CAF and adequacy are, therefore, needed to demonstrate a comprehensive and consistent picture of learners' writing performance influenced by cognitive task complexity.

Finally, language anxiety, as one factor of individual differences, may interfere with the effects of cognitive task complexity on L2 performance (Robinson, 2011a). The role that anxiety plays in the effects of cognitive task complexity on EFL performance is inconsistent in the limited research, especially in the writing field, with the small set of studies undertaken all correlational (MacIntyre & Gardner, 1994; Révész, 2011; Robinson, 2007b; Trebits, 2016). Interviews were seldom used to "better understand the role of anxiety in language learning" and learners' "feelings about language learning" (Yan & Horwitz, 2008, p. 153). More needs to be understood about the impact of language anxiety on the effects of cognitive-task variability on EFL writing.

The research design and methodology of this study is described in the next chapter.

Chapter 3 Research Design and Methodology

3.1 Overview

This chapter starts with a description of the worldview adopted in the present study, followed by an overview of the research design and a brief introduction to the ethical considerations. Details of participants and instruments, data collected and the data analysis of the pilot study and each phase of the main study are then presented.

3.2 Worldview

The pragmatic worldview is adopted as the philosophical underpinning for the methodological approach for this research. The pragmatic worldview originates from “actions, situations, and consequences rather than antecedent conditions” (Creswell, 2013, p. 10). Pragmatic researchers focus on the research problem and use all approaches available to address the problem (Creswell, 2013). Based on the pragmatic worldview, mixed-methods strategies of inquiry are best suited for this study, as the strategies help “to achieve an elaborate and comprehensive understanding of a complex matter, looking at it from different angles” (Dörnyei, 2007, p. 164). The pragmatic worldview and mixed-methods strategies are appropriate to be employed in this study to address the different aspects of the research questions (RQ) on the complexities of the relationships among cognitive-task variability, planning type, language anxiety and EFL writing.

3.3 Overview of the Research Design

The present study has three objectives: (a) to validate the designed cognitive-task variability; (b) to examine the potential isolated and synergistic effects of cognitive-task variability and planning type on EFL writing; and (c) to explore the relationships among L2 writing anxiety, cognitive-task variability, planning type and EFL writing production. Mixed methods include dual-task methodology, self-report questionnaires, writing tasks and interviews in order to achieve a full understanding of the complex target issue and corroborate findings through triangulation (Dörnyei, 2007).

The main study used mixed methods in which a predominantly quantitative phase was followed by a qualitative phase. Phase I of the main study was designed to address RQ1 to validate the manipulations of cognitive-task variability in each writing task by using a quantitative method. Phase II of the main study included two quantitative stages (Stage I and Stage II) and a qualitative stage (Stage III). Stage I of Phase II aimed to answer RQ2 to evaluate the potential isolated and synergistic effects, of cognitive-task variability and planning type on EFL writing. Stage II and Stage III of Phase II were designed to explore the overarching RQ3 which investigates the relationships among anxiety, cognitive task complexity and writing production but from quantitative and qualitative perspectives, respectively.

Altogether 292 second-year undergraduate English majors from China were recruited for the present study on a voluntary basis. Two dual-task experiments, two self-reported questionnaires and two writing tasks with different cognitive demands were the data sources for the quantitative stages (Phase I and the two quantitative stages of Phase II). The qualitative stage of Phase II contributed to an in-depth understanding of the quantitative findings of Stage II of Phase II and, probed what writers with different anxiety levels actually did and felt when engaged in different conditions. Qualitative data were collected via an in-depth semi-structured interview. Table 3.1 presents an overview of the research design.

Table 3.1 An Overview of the Research Design

Phases	Research questions	Instruments	Participants
Pilot Study	Pilot instruments	• Dual-task method design • Writing tasks	N = 62
Phase I	**RQ1**: *What are the relationships between presumed cognitive-task variability and cognitive load? If relationships exist,* a) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by self-perceived task difficulty? b) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by dual-task method?	• Questionnaire (TDMEQ) • Dual-task experiments	N = 91
Phase II-Quantitative stages	***Phase II-Stage I*-RQ2**: *What are the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing?* a) What is the potential effect of cognitive-task variability manipulated by varying the number of argument elements and reasoning demands on Chinese EFL students' argumentative writing? b) What is the potential effect of planning type (i.e., pre-task planning and online planning) on Chinese EFL students' argumentative writing? c) What are the potential synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing?	• Writing tasks	N = 201
	***Phase II-Stage II*-RQ3**: *What are the relationships among students' perceived anxiety, cognitive-task variability, planning type and EFL writing production?* a) What is the level of second language writing anxiety reported by university students in EFL writing?	• Questionnaire (SLWAI) • Writing tasks	N=108 (selected from Stage I of Phase II)

	b) What are the potential effects of second language anxiety, cognitive-task variability and planning type on students' EFL writing?		
Phase II-Qualitative stage	***Phase II-Stage III*-RQ3**: *What are the relationships among students' perceived anxiety, cognitive-task variability, planning type and EFL writing production?* c) What are students' perceptions of second language writing anxiety related to cognitive-task variability? d) What are students' perceptions of second language writing anxiety related to planning type?	• Semi-structured interview	N=15 (selected from Stage II of Phase II)

Note. SLWAI = Second Language Writing Anxiety Inventory; TDMEQ = Self-perceived Task Difficulty and Mental Effort Questionnaire.

3.3.1 Participants

A total of 292 second-year English majors were successfully recruited on a voluntary basis in two rounds from undergraduate programmes of two universities in Shandong province, China. Convenience sampling was used, as the universities were geographically proximate and easily accessible to the researcher. A shortcoming of convenience sampling, unrepresentativeness (Dörnyei, 2010), is overcome to a large extent in this research, because the undergraduate programmes for English majors in the universities in China were developed based on the guidance of *The English Teaching Syllabus for English Majors*, which provides specific and detailed requirements for English-major students (see more details in Section 1.1). English majors in Chinese universities, therefore, usually constituted a fairly homogeneous group in terms of educational background.

The curriculum designs and programmes of the two universities for English majors were similar. English writing was a compulsory course offered consecutively in the first two years in these two universities. The second-year English majors were chosen because second-year English majors from these two universities had almost finished the required writing courses according to the universities' teaching plan when the researcher initiated data collection. At this stage, it could be expected that the students had grasped the basic knowledge of writing and reached an intermediate level of English proficiency. All English majors in their first two years in the two universities had similar compulsory courses designed to develop their English competence in reading, writing, listening and speaking. Students in their third and fourth year took elective courses in their major, such as linguistics, literature, and business English. By choosing second-year English majors, who took the same required courses, some potentially confounding variables were controlled.

All participants were Chinese students with Chinese the language in which they were most proficient. Chinese is the commonly used national language and required in the national curriculum. Their ages ranged from 19 to 22. The participants had an average of 11.27 years of English learning at the time of this study, and none had overseas learning experiences. Participants were similar in age, educational background, English learning history, and English proficiency. Table 3.2 presents an overview of participant information; detailed information for each stage can be respectively found in section 3.5.1, 3.6.1, and 3.7.

In Phase I of the main study, 91 English majors from one university were randomly assigned into three experimental conditions: 31 in the single-task condition, 30 in the dual-task Condition 1 and 30 in the dual-task Condition 2. In Stage I of Phase II, another 201 participants from two universities were randomly assigned into two experimental groups, pre-task planning group and online planning group, and one comparison group. Based on students' responses to the SLWAI, participants in the three groups were classified into high- and low-anxiety levels to analyse further the relationships among language writing anxiety, cognitive-task variability, planning type and students' writing performance in Stage II of Phase II. In Stage III of Phase II, 15 students, based on purposive sampling from the high- and low-anxiety groups, were invited to participate in individual interviews.

Table 3.2 An Overview of Participant Information

Phases	Objectives	Total	Participants	Group information
Phase I	Validate the manipulation of cognitive-task variability by using multiple independent measures	91	Second-year English majors	• Single-task condition (N = 31) • Dual-task experimental Group 1 (*di&du*) (N = 30) • Dual-task experimental Group 2 (*di*) (N = 30)
Phase II-Quantitative stage	**Stage I**: Investigate the isolated and synergistic effect of cognitive task complexity on EFL writing production	201 ↓	Second-year English majors	• Experimental group: ➢ PTP (N = 68) ➢ OLP (N = 68) • Comparison group (N = 65)
	Stage II: Investigate the relationships among writing anxiety, cognitive-task variability, planning type and EFL writing	108 ↓	Second-year English majors	• PTP-H (N=18), PTP-L (N=18) • OLP-H (N=20), OLP-L (N=16) • CG-H (N=18), CG-L (N=18)
Phase II-Qualitative stage	**Stage III:** Investigate the relationships among writing anxiety, cognitive-task variability, planning type and EFL writing ➢ Secure an in-depth understanding of the quantitative data in Stage II of Phase II ➢ Probe more deeply what writers with different anxiety levels actually do and feel under different conditions	15	Second-year English majors	• 3 high-anxiety student from PTP; 2 low-anxiety from PTP • 3 high-anxiety student from OLP; 2 low-anxiety from OLP • 2 high-anxiety student from CG; 3 low-anxiety from CG

Note. PTP=Pre-task planning, OLP=Online planning, CG=Comparison group, H=High-anxiety level, L=Low-anxiety level.

3.3.2 Instruments

This section provides an overview of the instruments used in the present research, namely, self-report questionnaires, writing tasks, interviews and dual-task method experiments. It also presents the rationales for selecting these measures.

3.3.2.1 Self-report questionnaires

Two well established self-report questionnaires, SLWAI and *Self-perceived Task Difficulty and Mental Effort Questionnaire* (TDMEQ) were used to evaluate participants' L2 writing anxiety and perceptions of task difficulty and mental effort. Self-report questionnaires were chosen as the most appropriate approach for the present study, because they can effectively yield factual, behavioural and attitudinal data, as well as, successfully gather data in large-scale research (Dörnyei, 2010). Self-report questionnaires, with questions on perceptions of anxiety, task difficulty and mental effort were used to investigate the current situation of students' L2 writing anxiety and their perceptions of task difficulty and mental effort.

3.3.2.2 Writing tasks

Three argumentative essays were employed in the present study; one was used as the pre-test to evaluate participants' English-writing proficiency, and the other two, with the same topic but different elements and reasoning demands, were manipulated to present different levels, simple and complex, of cognitive-task variability. The argumentative genre was chosen in this research for three reasons. Firstly, argumentative writing is a popular and effective English-assessment tool to evaluate students' English writing ability, such as the writing section in IELTS and TOEFL, and CET and TEM (Teng & Zhang, 2016). Secondly, it has been one of the most used genres in academic writing and general courses at universities. Thirdly, the argumentative genre requires greater cognitive demands (Andrews, 1995; Freedman & Pringle, 1984), including more critical thinking and articulation of ideas, and more challenging use of linguistic elements (Hirose, 2003) than other genres. In the present study, the levels of cognitive demands were realized through varying the number of argument elements and reasoning demands in each of the two writing tasks.

3.3.2.3 Semi-structured interviews

A semi-structured interview is a compromise between a structured and an unstructured interview. It is a combination of "a set of pre-prepared guiding questions and prompts" and "the open-ended format" (Dörnyei, 2007, p. 136). It is an effective way to gain an in-depth understanding of participants' ideas and attitudes, and to check further the accuracy of information gained from other data-collection methods, such as close-ended questionnaires (Creswell, 2013; Dörnyei, 2007).

In this study, interviews were administered to the 15 students purposively chosen from different conditions in Phase II of the main study to provide an in-depth understanding of the quantitative data, and to probe further into what writers actually do and feel when they engage in different conditions.

3.3.2.4 Dual-task method

The dual-task method, a reliable and sensitive approach, has been used commonly in the field of experimental psychology to measure cognitive load in a direct and objective way (Brünken et al., 2003). As the name indicates, dual-task methodology involves performing a secondary task simultaneously with the primary task. In this study, dual-task methodology was the most appropriate approach to validate the writing-task manipulation. The primary tasks were two writing tasks designed to have different cognitive loads; the secondary task was a simple auditory stimulus.

3.3.3 Procedures for data collection

Data were collected in three stages (see Figure 3.1). Before conducting the experiment, a study was conducted to pilot the instruments and establish the time limits for the writing tasks (see Section 3.4). At the beginning of each stage, participants signed consent forms according to the regulations of the University of Auckland Human Participants Ethics committee, after being informed of the ethical issues associated with participating in the study as well as the research design, following which they completed the background questionnaire.

In Stage 1, Phase I of the main study was completed, and at the same time the pre-writing test in Phase II of the main study was implemented. Ninety-one students in three groups first completed a writing task with dual-task or TDMEQ. Another writing task was then completed after two weeks to avoid a similar topic effect from the former writing. Two writing tasks were counterbalanced to avoid potential practice and fatigue effects; half the participants took the simple task first, and half took the complex one first. In the meantime, 201 participants (see Section 3.6.1) were asked to finish the pre-writing test, which was designed to obtain a reliable and valid writing proficiency test score of all the participants. Based on the results of the pre-tests, participants in Phase II were matched to the CG, PTP and OLP groups (which are described in detail in Section 3.6.1)

When the cognitive-task variability had been validated with the intended levels of cognitive load, the quantitative stages of Phase II were conducted (Stage 2 in Figure 3.1). Students first completed SLWAI, and then the students under three planning conditions (following different writing instructions) were asked to complete two writing tasks at a one-month interval in two cycles. The two writing tasks, simple and complex, were counterbalanced. Based on participants' responses to the SLWAI, participants in the three groups were classified as having high or low anxiety using Cluster Analysis. At Stage 3, the qualitative stage of Phase II of the main study was conducted, where 15 participants were purposively recruited to participate in the semi-structured interview.

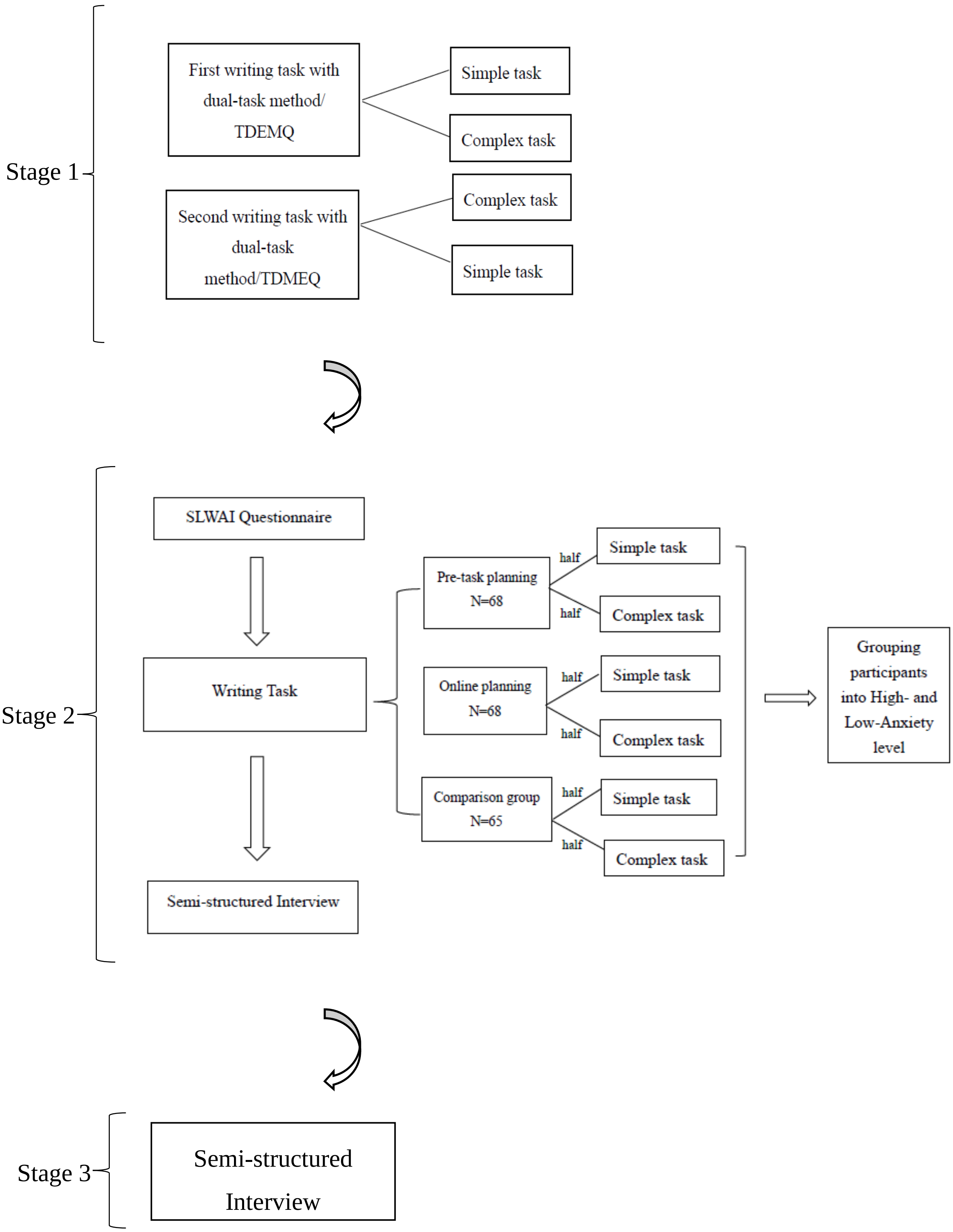

Figure 3.1 Design and procedure for the main study

Three operating stages in the main study. TDMEQ=Task Difficulty and Mental Effort Questionnaire; SLWAI=Second Language Writing Anxiety Inventory.

3.4 Pilot Study

A total of 62 students who shared the same context with the participants in the main study were recruited on a voluntary basis. The pilot study was conducted on both the dual-task experiments and the writing tasks.

3.4.1 Pilot the dual-task experiments

In this section, two dual-task experiments with different versions of the secondary task were piloted with two purposes. First, to evaluate whether the design of dual-task experiments was reasonable (i.e., frequency of the auditory stimulus was neither too low to reflect participants' cognitive loads of the primary task, nor too high for participants to concentrate on the primary task). Second, to examine the clarity and operability of the prompts and procedure of the dual-task experiments.

In the two dual-task experiments, the primary task was the simple and complex writing tasks in the main study with the topic of "choosing the best roommates". It required participants to choose one or two pairs of the best roommates out of four or six candidates, each marked by four or six properties. Participants were required to complete the primary writing tasks with at least 250 words within 40 minutes. Participants wrote the simple task first and the complex second. The two tasks were completed with a two-week interval. Two auditory versions of secondary task were designed, based on previous studies (Brünken et al., 2003; Marcus et al., 1996; Reed et al., 1985; Révész, 2014; Révész et al., 2014; Sasayama, 2016; Schoor et al., 2012).

3.4.1.1 Pilot dual-task Experiment 1

Ten participants were recruited voluntarily to pilot the dual-task Experiment 1. Participants conducted the experiment individually using the researcher's personal computer.

In Experiment 1, the design of secondary task with the two-sound stimulus was inspired by previous studies (Brünken et al., 2003; Marcus et al., 1996; Reed et al., 1985; Révész, 2014; Révész et al., 2014; Sasayama, 2016; Schoor et al., 2012). The computer was programmed to randomly generate a high pitch sound and a low pitch sound—*di* (1000Hz) and *du* (300Hz) for 150ms in a variable interval between 30 and 60s.

Half the participants were asked to react as fast and accurately as possible to *di* and ignore the sound *du* when they were writing the primary task. Another half were asked to react to *du*, while ignoring *di*. Reaction time and accuracy were recorded by E-prime 2.0.

Prior to the pilot experiment, there was a practice phase in which participants were instructed to perform the secondary task with writing a self-introduction as the primary task to get familiar with the dual-task procedure, the two sounds and E-prime environment. In the practice phase, participants first read task instructions presented in Chinese and were encouraged to ask any questions about the instructions, prompts and procedures. Following confirmation of their understanding of instructions, participants were asked only to carry out the secondary task without the primary task to collect the baseline reaction time and accuracy rates. In the final stage, participants simultaneously performed the primary writing task and secondary task. During the pilot experiment, the researcher took observation notes to record any questions or unexpected interruptions to the procedure.

In their feedback, participants reported that the instructions were understandable, the two sounds in the pilot Experiment 1 was easily recognized, and the frequency of the stimulus was reasonable. However, they did not know how to close the software, E-prime, when they finished writing the primary task. A small amendment, therefore, was made that at the very beginning of the experiment, participants would be taught how to close the E-prime when it was running. Students' reaction time to the experimental condition was longer than to the baseline condition, and the accuracy in the baseline condition was much higher than the experimental condition. These results confirmed the validity of the dual-task experimental design.

Finally, most of the participants in the pilot finished writing within 40 minutes, as they were performing the secondary task. Forty minutes, therefore, was set for the dual-task in Experiment 1.

3.4.1.2 Pilot dual-task Experiment 2

Another 10 students, recruited voluntarily to pilot Experiment 2, completed the experiment individually by using the researcher's personal computer.

For the secondary task, based on Reed, Burton and Kelly (1985), Experiment 2 adopted a simple continuous auditory task with one sound stimulus. A sound *di* with 1000Hz pitch was generated at random intervals in the range of 15-30s (Brünken, Steinbacher, Plass, & Leutner, 2002). Participants were instructed to respond as fast and accurately as possible to the stimulus on hearing the sound. Reaction time and accuracy were obtained by E-prime 2.0. The warm-up practice and baseline test were conducted before the experiment, and the researcher took observation notes throughout the procedure.

In contrast to expectation, seven participants (70%) found the frequency of the sound stimulus was too high for them to concentrate on completing the primary writing task. Some amendments, according to students' feedback and previous research (Madrid, Van Oostendorp, & Melguizo, 2009), were made to the frequency of the stimulus, with the range of the intervals changed to 20-40s.

Another 10 participants were asked to pilot the new version of Experiment 2. This time, the frequency of sound stimulus with *di* presented at random intervals in the range of 20-40s, when they were simultaneously carrying out primary writing tasks, was found acceptable by participants. Therefore, the 20-40s frequency of the secondary task was used in the present study. Additionally, most of the participants finished the whole process of dual-task Experiment 2 within 40 minutes. Forty minutes, therefore, was set for the dual-task Experiment 2. Students' reaction time to the experimental condition was found much longer than the baseline condition, and the accuracy of the baseline condition was higher than the experimental condition, which confirms the validity of the dual-task experimental design.

3.4.2 Pilot the two writing tasks

In this section, the two experimental writing tasks were piloted: (a) to pre-evaluate the manipulation of the cognitive-task variability by obtaining experts' and students' feedback on the two writing tasks and reasons for their feedback; (b) to check whether the writing task instructions were clear to students, whether the writing prompts were linguistically and culturally understandable to student, and whether these selected argumentative writing topics were writable for students; and (c) to determine the allocation of the planning and writing time under different planning conditions.

3.4.2.1 Pre-evaluation of the manipulation of cognitive-task variability

The two writing tasks with the topic of "choosing the best roommates" were piloted before conducting the main study. An experienced Chinese teacher was invited to edit the translated versions to make the instructions and prompts clear and easy to understand.

Seven Chinese PhD students from the University of Auckland, who focused on second language writing research, were invited to voluntarily judge the cognitive load level of the two writing tasks, and give feedback on the instructions and prompts. They were asked to respond to an open-ended questionnaire to find which task had greater cognitive load and was more difficult (Task 1 was designed-to-be simple and Task 2 designed-to-be complex) and explain why they made this choice. As a result, all seven experts found that Task 2 was more difficult than Task 1. They thought that Task 2 included more elements, consumed more mental effort, needed more critical thinking and required more challenging use of linguistic elements, as the following excerpts illustrate:

> "Students will take more things into consideration and make more comparisons in Task 2."
>
> "More information and more choices are included in Task 2, which could increase the difficulty for students to finish the essay."
>
> "Task 2 includes more two elements, and students also should make two more decisions in Task 2 than Task 1. For Task 2, it is hard to make decisions because the choices are closely related and there are no perfect pairs for the best roommates. Task 2 costs more cognitive load."
>
> "It takes more cognitive demands to make a decision in Task 2. The topic is interesting, which makes me think of my college life."
>
> "Compared with Task 1, in Task 2, students will make more critical thinking before writing, and need more linguistic expressions and vocabulary to explain the reasons during the writing."

In feedback on prompts, two of the experts suggested that the researcher replace the choice "cooking" in the column of *Hobbies*. This is because in China cooking is forbidden in the school dormitory and most high school students are unfamiliar with the cooking topic, unlike students in Western countries. Based on their advice, "cooking" was replaced by an alternative choice “playing football”, which is more suitable in the Chinese context.

In addition, the two writing tasks were piloted with five participants. They were encouraged to raise any questions and report difficulties they found when they read the instructions and prompts of the two writing tasks. They were also asked to evaluate how much cognitive load they would experience if they were asked to complete the two writing tasks. According to participants’ feedback, the instructions and prompts of the writing tasks were clear to them. Similarly, all the participants found Task 2 would consume a greater cognitive demand and was more difficult. The final version of the two writing tasks is in Appendix C.

3.4.2.2 Determine the planning and writing time under different conditions

Another 27 participants in the pilot study were recruited to voluntarily complete the two writing tasks. Fifteen were asked to write the simple task, and the remaining 12 students performed the complex task. Prior to their writing, participants were told that their task performance would be assessed based on content, organisation, language, vocabulary and mechanics, and they were not allowed to consult any books, dictionaries or other materials during writing. In this pilot study, no time limit was set, and participants’ completion time was noted.

Piloting results indicated that the slowest participants spent 35 and 37 minutes completing the simple and complex tasks respectively, and the fastest writers finished the simple task in 24 minutes and the complex task in 26 minutes. To ensure all participants could successfully complete the tasks, a decision, following the instructions of Ellis and Yuan (2004), was made to allocate times for the different experimental conditions in the main study. In the pre-task planning condition, 15 minutes was given for planning before writing, as according to previous studies (Ellis & Yuan, 2004; Foster & Skehan, 1996; Mehnert, 1998; Ong & Zhang, 2010), at least 10 minutes planning time would provide measurable effects. Twenty-five minutes was given to complete writing the tasks, because in 25 minutes, participants would be “pressured to perform the task with limited opportunities for online planning” (Ellis & Yuan, 2004, p. 69). In the online planning condition, participants had 40 minutes to write to provide them ample time for online planning while they are writing. In the comparison group, participants were told to write the task in their usual style within 40 minutes. According to the 27 participants' writing performance, they can finish the task by writing an argumentative genre, and most of them can meet the 250 words requirement. More detailed instructions of the three experimental conditions can be found in Section 3.6.3.

3.4.3 Modifications for the main study

The trial procedure, throughout this pilot session, validated some of the methodological propositions. The experimental conditions were not fully tested because of the complexity of the research design which would have required a large sample of students in the pilot study. Nevertheless, the pilot study led to some modifications of the original methods and data-collecting instruments. The following changes were made for the main study.

1. Instruments: Instructions for the prompts of the writing tasks
   1) For the two experimental writing tasks prompts, "cooking" in the column of *Hobbies* was replaced by "playing football" to suit the Chinese context.
2. Procedure: Design of dual-task methodology, and time allocation under different conditions for Phase II of the main study
   1) At the very beginning of the dual-task experiments, participants would be taught how to close the E-prime when it was running and the data would be automatically saved by the software.
   2) In the secondary task, the frequency of the stimulus of the second version of the secondary task was amended, with the range of the intervals changed to 20-40s.
   3) For the completion of the dual-task experiments, 40 minutes was set for the two dual-task experiments.
   4) For the allocation of the planning and writing time, under the PTP condition, 15 minutes was set for pre-task planning and 25 minutes to write; under the OLP condition, participants would be instructed to write immediately for 40 minutes and to plan during their writing; in the comparison group, participants would be told to write the task in a natural way within 40 minutes.

3.5 Phase I—Validation of Cognitive-task Variability

The purpose of Phase I of the main study is to validate the designed cognitive-task variability within the two experimental writing tasks by using multiple independent measures. Phase I is guided by the following research questions:

What are the relationships between presumed cognitive-task variability and cognitive load? If relationships exist,

(a) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by self-perceived task difficulty and mental effort self-assessment?

(b) Do different intended levels of cognitive-task variability result in different levels of cognitive load during task performance, as measured by dual-task method?

3.5.1 Participants

Ninety-one second-year English majors in three intact classes from a university in China were voluntarily recruited in Phase I (92 students were initially recruited; 1 was removed as an outlier in the dual-task Experiment 2). They shared the same context as the participants in Phase II of the main study. The mean age was 20 with a range of 19-22, and most of the participants were female (90% female and 10% male). No participants had overseas learning experiences. Their years of learning English ranged from 8 to 14 years with an average of 10.88 years. No significant differences were found in students' English proficiency across these three intact classes in their last semester final exam results of *Integrated English Course* and *English Writing Course*. Students in three intact classes as a whole were, therefore, randomly assigned into three groups with different experimental conditions. Group 1 with 31 students was the single-task condition group in which participants were asked to complete the writing tasks as well as the following TDMEQ. Group 2 with 30 students and Group 3 with 31 students were the dual-task condition groups in which participants were required to perform the writing tasks as the primary task and secondary task simultaneously.

3.5.2 Instruments

The instruments used in Phase I of the main study included a questionnaire eliciting self-perceived task difficulty and mental effort, two writing tasks as the primary task, and two versions of auditory tasks as the secondary tasks.

3.5.2.1 Self-perceived task difficulty and mental effort questionnaire

The TDMEQ was adopted from Révész et al. (2016). The paper-based self-rating scale, translated into Chinese, included six items with a 9-point Likert response scale ranging from 1 (very easy/not at all) to 9 (very difficult/intensively). Cronbach's Alpha showed the questionnaire to reach acceptable reliability α =. 75. Participants in Group 1, the single-task condition completed this scale when they finished each writing task. Based on the precedent of Révész et al. (2016), only two items (Item 1 and Item 3) out of six, relevant to measuring cognitive load, are discussed in this study. The two items were as follows:

(Item 1) This task required no mental effort at all	1	2	3	4	5	6	7	8	9	This task required extreme mental effort.
(Item 3) This task was very easy.	1	2	3	4	5	6	7	8	9	This task was very difficult.

Item 1 assessed participants' perceptions of mental effort and item 3 identified participants' feelings of the task difficulty. The questionnaire can be found in Appendix B.

3.5.2.2 Two writing tasks

The two writing tasks were adapted from Cho's (2015) work with the topic "choosing the best roommates". The prompts of the two writing tasks were translated into Chinese to ensure all the participants could fully and quickly understand the topics and instructions. Based on the pilot study (see Section 3.4.2.1), the prompts and instructions were modified to suit the EFL learning context in China. These tasks were chosen because students living in the dormitory on the school campus in China were familiar with, and interested, in this topic.

In the two argumentative writing tasks, cognitive-task variability was manipulated by varying the number of argument elements and different reasoning demands. The simple task required learners to choose one pair of best roommates from four candidates, each marked by four properties: hobbies, personality, studying style, and sleeping patterns (see Appendix C). The complex task asked learners to choose two pairs of best roommates from six candidates, each marked by six properties: hobbies, personality, studying style, sleeping pattern, favourite subjects and individual sanitary habits (see Appendix C). The complex task, compared to the simple version, is characterized by more argument elements and greater reasoning demands, with two more candidates and two more properties of each candidate from which to choose. As a result, the complex task was expected to impose greater cognitive demands on students. The rationale for this manipulation is that to complete the tasks with a greater number of argument elements and more reasoning demands is more cognitively demanding than those with fewer elements and demands on reasoning (Cho, 2015; Kuiken & Vedder, 2007a; Robinson, 2005, 2007b).

3.5.2.3 Secondary tasks in dual-task experiment

Two versions of simple, continuous auditory tasks designed as secondary tasks were carefully revised based on the pilot studies (see Section 3.4.1).

In dual-task Experiment 1, Secondary Task 1 with a two-sound stimulus, *di* and *du*, was conducted by participants in Group 2, and in dual-task Experiment 2, Secondary Task 2 with a one-sound stimulus, *di*, was performed in Group 3.

Secondary Task 1. Stimuli were two sounds—a high pitch sound *di* (1000Hz) and a low pitch sound *du* (300Hz). The computer was programmed to generate *di* and *du* for 150ms at random intervals in the range of 30-60s. Participants were instructed to press “Z” in the keyboard as fast and accurately as possible, when they detected the sound through an earphone to which they were required to respond. Half the participants were required to react to *di* ignoring *du*, and half to *du* ignoring *di*. The reaction time, error and accuracy rates were obtained from E-prime 2.0 in milliseconds.

Secondary Task 2. Using E-prime, the sound *di* (1000Hz) was programmed to occur for 150ms in a variable interval between 20 and 40s. Participants, upon hearing *di*, were asked to react as fast and accurately as possible to the stimulus. Participant’s reaction time and accuracy in responses to *di* by pressing "Z" in the keyboard were recorded by E-prime 2.0 in milliseconds.

3.5.3 Procedures

Participants in Group 1, the single-task condition, were asked to complete two writing tasks, each writing task followed by the TDMEQ. Two-week intervals occurred between the two writing tasks, and two writing tasks were counterbalanced to avoid practice and fatigue effects. The schedule is summarised in the following table (Table 3.3):

Table 3.3 Procedure for Single-task Condition

Group 1	Week one	Week four
31 second-year English majors	First writing task (40 min) + TDMEQ (5 min)	Second writing task (40 min) + TDMEQ (5 min)

Participants, in Group 2 and Group 3, the dual-task condition, were required to complete the primary writing tasks and secondary tasks simultaneously. Participants were tested individually in groups of up to 16 students at one time in the school computer room. Each experiment was conducted twice by the researcher with the help of two research assistants. The first time took about 90 minutes for each participant, including a practice phase, a baseline test and the first dual-task test, and the second time took about 60 minutes, including a warm-up phase and the second dual-task test. There was a two-week interval between the two experiments, and the two primary writing tasks were counterbalanced. Task instructions were presented in Chinese to ensure participants can easily and fully understand them. The procedures are summarised in the following table (Table 3.4):

Table 3.4 Procedure for Dual-task Experiments

Group Time	Group 1 (n = 30) (*di* & *du* secondary task version)	Group 2 (n = 31) (*di* secondary task version)
First time (Week one)	Practice phase (15 min) ↓ Rest (5 min) ↓ Baseline test (20 min) ↓ Rest (10 min) ↓ First dual-task test (40 min)	Practice phase (15 min) ↓ Rest (5 min) ↓ Baseline test (20 min) ↓ Rest (10 min) ↓ First dual-task test (40 min)
Second time (Week four)	Warm-up phase (15 min) ↓ Rest (5 min) ↓ Second dual-task test (40 min)	Warm-up phase (15 min) ↓ Rest (5 min) ↓ Second dual-task test (40 min)

The practice phase was expected to help participants to get familiar with task instructions and the experimental procedures in the dual-task experiment, in which participants were asked to write a piece of self-introduction on the paper and, at the same time, react to the secondary task. During this phase, participants were encouraged to ask the researcher and research assistants any questions they had regarding the procedures, and the researcher, at the same time, monitored them to ensure that they were appropriately following the task instructions. Some students were given additional instructions as needed. After this phase, participants had a 5-minute rest.

The participants were then informed to complete the secondary task for 20 minutes without an added primary writing task. They were told to respond to the sound stimulus as fast and accurately as possible, and their reaction time and accuracy rate would be recorded. After this phase, participants had a 10-minute rest.

The participants next moved on to the actual dual-task experiment. Before each task, participants had three minutes to understand the prompts and instructions of the writing tasks. They were informed that the writing tasks were a test of their writing ability and would be rated in terms of content, organisation, vocabulary, language and mechanics, and they were not allowed to consult any papers, books, dictionaries or other materials during the experiment; and that their reaction time and accuracy rate would be recorded. The participants then started the actual dual-task experiment; this phase took 40 minutes.

After a two-week interval, participants took part in the second experiment. After the participants completed a 15-minute warm-up exercise to become acquainted with the dual-task procedure, they engaged in the actual dual-task experiment for 40 minutes.

3.5.4 Data analysis

Quantitative data collected via TDEMQ and the dual-task method were first input into SPSS 24 and cleaned. The normal distribution, missing values and outliers were examined.

3.5.4.1 Analysis of self-perceived task difficulty and mental effort questionnaire

Independent variables were the two levels of difficulty of the two writing tasks (i.e., the simple and complex tasks), while the dependent variables were participants' perceived task difficulty and mental effort presented via Item 1 and Item 3 of the TDMEQ. The normal distribution of dependent variables was first examined, and no missing values and no outliers were found in the data.

The paired-samples *t*-test compares two means, which come from the same entities (Field, 2009; Pallant, 2013). Therefore, the paired-samples *t*-test was applied to investigate whether there were significant differences in Item 1 and Item 3 of TDMEQ tested at two different times (following two writing tasks).

3.5.4.2 Analysis of the dual-task results

Participants' responses to the secondary, sound-detection task were recorded every time they pressed the "Z" in keyboard to indicate their reaction to the sound stimulus. Their responses were analysed in three ways: (a) reaction time of the correct responses, (b) accuracy rate, following previous research (Révész et al., 2016), and (c) two types of error rates, based on previous research (Sasayama, Malicka, & Norris, 2015). In this section, the independent variables were the three task conditions (i.e., no primary task under the baseline condition, the simple primary task under the dual-task condition, and the complex primary task under the dual-task condition), and the dependent variables were participants' reaction time, accuracy rate, and error rates in secondary tasks.

For the Secondary Task 1 with *di* and *du* version (in dual-task Experiment 1), accuracy was calculated by dividing the number of times participants correctly reacted to *di* and ignored *du* (half of them correctly reacted to *du* and ignored *di*) by the total number of stimuli (both *di* and *du*) occurring. Error I was calculated by dividing the total number of instances where a participant missed the sound to which they were required to react by the total number of stimuli (both *di* and *du*). Error II was calculated by dividing the number of instances where a participant erroneously reacted (participants reacted to the sound which is not required) by the total number of occurrences of the sound that should be ignored by participants. Reaction time was considered for only correct responses.

As for Secondary Task 2 with *di* version (in dual-task Experiment 2), reaction time and accuracy rate were calculated. Accuracy rate was calculated by dividing the number of times participants correctly reacted to *di* by the total number of occurrences of *di*. Reaction time was calculated by dividing the summed reaction time of accurate responses to *di* by the total number of correct responses to *di*.

All data were first inspected for missing values and outliers. No missing values and no outliers were found in the data of Group 2. In Group 3, one participant's data were excluded as an outlier due to a cut-off set at three standard deviations from the mean.

To examine the effects of cognitive-task variability on the dual-task data, one-way repeated-measures ANOVAs were conducted. Repeated-measures ANOVA compares several means when those means came from the same entities, and one-way repeated-measures ANOVA includes one factor with at least two levels, and those levels are dependent (Field, 2009; Pallant, 2013). In this section, the factors were reaction time, accuracy rate and error rates, and the three levels were the baseline condition, simple task as the primary task in the dual-task condition, and complex task as the primary task in the dual-task condition. A series of one-way repeated-measures ANOVAs, therefore, was conducted to test whether there were significant differences in reaction time, accuracy rate and error rates across three cognitive-task variability levels. Assumptions of one-way repeated-measures ANOVA were first examined, including relevant boxplots and scatterplot matrixes for normality and Mauchly's test.

3.5.4.3 Effect size

Effect size, a measure of the strength of an effect, can help to ascertain the practical significance of statistical significance (Cohen, 1988). There are different types of effect sizes, and this study reports the partial eta-squared (η_p^2) for ANOVA (Mixed ANOVA, one-way ANOVA and one-way repeated-measures ANOVA), and Cohen's *d* for *t*-test (paired samples *t*-test and independent samples *t*-test). Table 3.5 shows the magnitude of the effect sizes for different statistics.

Table 3.5 Summary of Magnitude of Effect Sizes (Cohen, 1988)

Data analysis	Effect size	Small	Medium	Large
t-test	Cohen's *d*	.20	.50	.80
ANOVA	η^2	.01	.06	.14

3.6 Phase II—Quantitative Stage of the Main Study

3.6.1 Participants

Altogether 201 English majors in their second year were recruited on a voluntary basis (209 students were initially recruited and eight were removed due to missing data or as outliers). These students came from seven intact classes from two universities in China (three classes from one university and four classes from another university). No significant differences in the students' English proficiency and English writing proficiency were found between intact classes of each university in the semester final examination of both *Integrated English Course* and *English Writing Course* in the two universities. Students in three intact classes from University 1 as a whole were randomly assigned into the CG, PTP and OLP groups. Students in the first three intact classes (Class 1, Class 2, and Class 3) from University 2 served as three groups labelled as PTP, OLP and CG. Students from Class 4 in University 2 were categorized into high (n = 9), average (n = 12) and low (n = 9) levels of writing proficiency according to their last semester final examination scores. Then, 3 low-, 4 medium- and 3 high-proficiency writers were assigned to each of three sub-groups. The detailed information and students' average writing scores in the three sub-groups are shown in Table 3.6. No significant difference in the English writing proficiency of each sub-group was found, F (2, 27) = .025, p = .749.

Table 3.6 Average English Writing Score in Three Sub-groups for Class Four

	Sub-group 1	Sub-group 2	Sub-group 3
Low-proficiency	n=3	n=3	n=3
Medium-proficiency	n=4	n=4	n=4
High-proficiency	n=3	n=3	n=3
Total participants	n=10	n=10	n=10
Average writing score	79.4	79.6	79.1

The three sub-groups as a whole were, then, randomly assigned into the CG, PTP and OLP groups. In this way, a similar number of participants in each group was created (see Table 3.7), with 69 students in the PTP group, 71 in the OLP group and 69 in the CG. No significant differences were found in participants' *Writing Course* final examination, which indicated that general EFL writing proficiency levels across the three groups were about the same, as confirmed by their pre-writing test scores presented in Section 3.6.3.

Table 3.7 Participants under Three Conditions Initially Recruited in Phase II

	PTP	OLP	CG
Class 1-3 from University 1	29	31	31
Class 1-3 from University 2	30	30	28
Class 4 from University 2	10	10	10
Total	69	71	69

One student from the PTP group, two from the OLP and two from the CG missed either the pre-writing test or each of the two writing tasks during data collection. Data from five students were incomplete and, therefore, not used in the analysis. Furthermore, data of two participants from the CG and one from the OLP group were excluded as outliers with a cut-off set at three standard deviations from the mean. After excluding the extreme cases, data from 201 participants were coded and subjected to analysis.

The three intact classes in University 1 shared the same writing course lecturer and one lecturer also taught the English writing course for the four intact classes in University 2; the instructional programmes for English majors in their first two years in these two universities were quite similar. According to the background questionnaire, the participants were aged between 18 and 22, and all had received their previous English education in mainland China. The ratio of female to male was 87% to 13%. They had, on average, learned English for 11.44 years, but none of the participants had taken the IELTS tests. The demographic information of participants in each group condition is displayed in Table 3.8.

Table 3.8 Demographic Information of Participants in the Three Planning Groups

Group	Gender		Total number	Average Age	Average years of learning English	Overseas learning experience
	Male	Female				
PTP	9	59	68	20	11.65	None
OLP	12	56	68	20	11.15	None
CG	6	59	65	20	11.54	None

To address RQ3, participants in the three planning groups were classified further, based on their responses to the anxiety questionnaire, as a low or high level of anxiety in Stage II of Phase II. Twenty participants were associated to the online planning-high anxiety (OLP-H) group, 16 to the online planning-low anxiety (OLP-L) group, 18 to the pre-task planning-high anxiety (PTP-H) group, 18 to the pre-task planning-low anxiety (PTP-L) group, 18 to the comparison-high anxiety (CG-H) group and 18 to the comparison-low anxiety (CG-L) group.

3.6.2 Instruments

Two types of instruments, questionnaire and writing tasks, were employed in the quantitative phase of Phase II. The SLWAI was used to assess students' perceptions of L2 writing anxiety. A pre-writing test was used to confirm whether there were significant differences in writing proficiency among the three groups. Two writing tasks, manipulated with varying levels of cognitive-task variability, were employed to examine the potential effects of cognitive-task variability on students' EFL writing performance under the three planning conditions.

3.6.2.1 Second Language Writing Anxiety Inventory

The writing anxiety questionnaire in the present study was modified from Cheng's (2004) SLWAI to assess participants' level of L2 writing anxiety within the Chinese EFL context. This is a well-established self-report instrument with 22 items, numbered from 1 to 22 corresponding to three subscales—Somatic Anxiety, Avoidance Behaviour and Cognitive Anxiety. Seven of the items are negatively worded and require reverse scoring before being summed up to total scores, such as "While writing in English, I'm not nervous at all". In this study, the questionnaire is a 7-point Likert response scale ranging from 1 (strongly disagree) to 7 (strongly agree), and the total scores of these 22 items represent anxiety values ranging from 22 to 154. The higher the scores obtained, the more serious is the writing anxiety. The questionnaire was translated into Chinese. More information about the scale can be found in Table 3.9 as well as in Appendix A.

Table 3.9 An Overview of Second Language Writing Anxiety Inventory

Subscales	Total items	Item number	Example
Somatic Anxiety	7	2, 6, 8, 11, 13, 15, 19	I feel my heart pounding when I write English compositions under time constraint.
Avoidance Behaviour	7	4, 5, 10, 12, 16, 18, 22	I usually do my best to avoid writing English compositions.
Cognitive Anxiety	8	1, 3, 7, 9, 14, 17, 20, 21	While writing English compositions, I feel worried and uneasy if I know they will be evaluated.

Note. Adapted from "A measure of second language writing anxiety: Scale development and preliminary validation" by Y. S. Cheng, 2004, *Journal of Second Language Writing*.

3.6.2.2 Pre-writing test

Prior to the experiment, participants were asked to complete a pre-writing test to evaluate their English writing proficiency. The topic of the pre-writing test, chosen from IELTS, concerned whether universities should provide theoretical knowledge or practical training. Following is the writing topic and the prompts of the pre-writing test.

Some people think that universities should not provide theoretical knowledge, but to give practical training that is beneficial to society. Do you agree or disagree?

Give reasons for your answer and include any relevant examples from your own experience or knowledge. You should spend about 40 minutes on this argumentative writing task and write at least 250 words.

3.6.2.3 Two writing tasks

The two writing tasks shared the same topic, choosing the best roommates in a boarding high school. The two argumentative writing tasks were manipulated with different levels of cognitive-task variability by varying elements and reasoning demands, and were validated in Phase I (Section 3.5).

The simple version of the writing task required participants to choose two best roommates out of four candidates, while students were required to select four out of six candidates in the complex version. The writing topics were the same across the three groups, but the writing prompts differed with respect to the planning conditions. The following are examples of the simple version for three planning conditions. More detailed information can be found in Appendix C.

Imagine you can choose two best roommates from the four candidates when the new semester starts (two students in each dormitory). Four properties of candidates: hobbies, personality, studying style and sleeping patterns are listed in the table (omitted here) as follows.

Read the descriptions carefully and make a considered choice of the best roommates. Write an argumentative essay. Please provide at least 3 reasons to make your decision clear and convincing.

The best roommates you choose: _________ *&* __________

PTP: You have 15 minutes for planning and 25 minutes for writing. You can plan the content, organisation, and language, but your notes will be taken away before commencing of writing. Please write at least 250 words.

OLP: You have no time for pre-task planning but 40 minutes for writing and online planning. You should start writing immediately and plan while you are writing. Please write at least 250 words.

CG: You have 40 minutes on this task. Please write at least 250 words.

3.6.3 Procedures for data collection

Data for Phase II of the main study were collected from 2nd May 2018 to 10th July 2018. Prior to the start of data collection for Phase II, data for Phase I were collected and analysed. The results of Phase I showed that the manipulation of the cognitive-task variability of the two writing tasks was reasonable—the designed-to-be complex task yielded greater cognitive demands than the simple version. At the same time, all the 291 participants were asked to complete the pre-writing test within 40 minutes in one of the writing classes. The participants were then assigned to the CG, PTP and OLP groups (for the detailed allocation see Section 3.6.1). A one-way ANOVA test showed that no significant differences in the prewriting test were identified between each two groups, F (2, 198) = .554, p = .575.

At the beginning of Phase II, all participants in the three groups were invited to perform the SLWAI within 15 minutes after a writing course. Then, participants under the three groups were, each instructed to complete the two writing tasks at one-month intervals, and the two tasks were counterbalanced.

The three planning conditions were guided by the pilot study. Participants under the PTP condition, who were given 15 minutes for pre-task planning and 25 minutes to finish writing the task, were expected to be pressured to perform the task with limited opportunities for online planning. The instructions for their pre-task planning followed those of Ellis and Yuan (2004): Participants were provided with a sheet of paper to write notes during the planning process, but the notes were taken away before the task commenced to ensure that "the language elicited by all the tasks is produced within the specified time limit" (Ellis & Yuan, 2004, p. 70). Suggestions were made to the participants to plan their writing in terms of content, organisation and language. No other detailed guidance was provided.

In the OLP group, 40 minutes were allocated with no pre-task planning time given to participants prior to writing the tasks. Participants were asked to begin writing immediately and instructed to plan while they were writing. It was assumed that participants would have ample time for online planning while they were writing based on the pilot study described in Section 3.4.2.2.

In the comparison group, participants were requested to complete the task in their usual style within 40 minutes. The 40-minute limit is the same as a writing test in a traditional writing course. The task conditions are summarised in Table 3.10.

Table 3.10 Task Conditions

Task condition	Pre-task planning time	Writing time	Total time
PTP	15	25	40
OLP	0	40 (planning while writing)	40
CG	/	/	40

Note. "/" means no specific requirements.

In Stage II of Phase II, participants in the three groups were classified as high- and low-anxiety groups using Cluster Analysis based on participants' responses to the SLWAI, completed prior to the writing tasks.

3.6.4 Data analysis

Data analysis in Phase II of the main study involved two stages: Writing production analysis and statistical analysis. Analysing students' writing performance of the two tasks preceded the statistical analysis, and included coding the CAF and rating the adequacy. Then, the statistical analysis was conducted to address RQ 2 and RQ 3.

3.6.4.1 Writing production analysis

Multidimensional features of learners' writing performance were measured in terms of complexity, accuracy, fluency, and adequacy. Complexity included syntactic and lexical complexity. Adequacy, including content, organisation and overall writing quality, was evaluated based on Jacobs et al. 's (1981) rating scheme.

3.6.4.1.1 Syntactic complexity

Syntactic complexity was multi-dimensionally measured in four sub-constructs in the present study: overall complexity, complexity by subordination, and complexity via phrasal elaboration and complexity via coordination (Norris & Ortega, 2009) .

The overall complexity was measured by the mean length of T-unit (MLT). MLT is defined as the number of words divided by the number of T-units, which has consistently been used as a length-based measure (Frear & Bitchener, 2015; Ruiz-Funes, 2015; Yang et al., 2015; Yoon, 2017). A T-unit, proposed first by Hunt (1966) as “one main clause plus the subordinate clauses attached to or embedded within it” (p. 49), has been one of the most often used measures of L2 development (Wolfe-Quintero et al., 1998).

For local syntactic complexity, subordination was measured by clauses per T-unit (C/T) and dependent clauses per clause (DC/C) to capture the intermediate proficiency level complexity. Coordination, measured by T-units per sentence (T/S) is also included in the present study, since the participants in this study were all EFL writers, whose English proficiency was lower than that of native speakers. To capture the syntactic complexity development of advanced proficiency, mean length of clause (MLC) was used for phrase-level measures, coordinate phrases per clause (CP/C) for phrasal coordination, and complex nominals per clause (CN/C) for noun phrase complexity. In this way, students’ interlanguage development at different proficiency levels could be captured. These seven measurements were analysed by using the EFL SCA (Lu, 2010).

3.6.4.1.2 Lexical complexity

Lexical complexity in the present study was measured by lexical density and lexical diversity. Lexical density was analysed by using the EFL LCA (Lu, 2012). According to Lu (2012), lexical words were defined in the EFL LCA as “nouns, adjectives, verbs (excluding modal verbs, auxiliary verbs, ‘be’ and ‘have’), and adverbs with an adjective base, including those that can function as both an adjective and adverb (e.g., ‘fast’) and those formed by attaching the *–ly* suffix to an adjectival root (e.g., ‘particularly’)” (p. 192).

Lexical diversity was measured by WT/√2W which was calculated by taking text length into account (Carroll, 1967). WT/√2W stands for the total number of different word types divided by the square root of two times the total number of words. The number of words and word types was counted automatically by the EFL LCA (Lu, 2012).

3.6.4.1.3 Accuracy

Three general accuracy measurements, the number of errors per T-unit (Err/T), the number of errors per 100 words (EP100) and the ratio of error-free clauses (EFC/C), were employed to better examine learners' overall ability in using the second language.

EFC/C was chosen instead of error-free T-units, since several participants in this study made errors in every T-unit and in such cases the number of error-free T-units was zero. Additionally, the EP100 was calculated by dividing the total number of errors by the total number of words multiplied by 100. Err/T and EP100 were used to capture the information of the number and type of errors.

An error-free clause was defined as a clause that is free of any lexical, grammatical, or morphological errors. Errors were coded and tailed based on Polio and Shea's (2014) guidelines. An error was considered as any digression in syntax, morphology, and lexical choice, but not in punctuation or capitalization (Ellis & Yuan, 2004).

To sum up, Err/T and EP100 were employed to measure the presence of errors and EFC/C to capture the absence of errors, as Polio and Shea (2014) suggested that more than one measure of accuracy should be used in a study.

3.6.4.1.4 Fluency

Fluency, which refers to "the capacity to mobilize one's linguistic resources in the service of real-time communication" (Skehan, 1996a, p. 48), has been commonly used in previous studies (Ellis & Yuan, 2004; Kellogg, 1990; Ong & Zhang, 2010). Fluency in the present study is calculated by: the mean number of words produced per minute based on total time on task (Kellogg, 1990; Ong & Zhang, 2010).

3.6.4.1.5 Adequacy

Adequacy in terms of content, organisation and overall text quality was examined to explore whether participants met the pragmatic appropriateness of the tasks (Pallotti, 2009). Students' adequacy was scored based on the composition rating scheme of Jacobs et al. (1981). This marking rubric has been widely and successfully used in evaluating the writing proficiency levels of students in EFL programmes (Ong & Zhang, 2013; Teng & Zhang, 2018, 2020). The rubric, ranging from 34 to 100, used a weighted scoring scheme on a percentile scale. Five aspects of writing performance were measured: content (30%, 13-30), organisation (20%, 7-20), language (25%, 5-25), vocabulary (20%, 7-20), and mechanics (5%, 2-5). Each of the five components has four rating levels with clear descriptors of the writing proficiency for each level and a corresponding numerical scale. The separate scores for the content and organisation subcategories of learners' writing production and students' overall writing score were analysed. Language, vocabulary and mechanics were not discussed in the present study, since they, to some extent, overlapped the indexes of syntactic complexity, lexical complexity and accuracy. See Appendix D for details of the writing scoring scheme.

3.6.4.2 Statistical analysis

Data collected in the quantitative stage of Phase II via the writing tests and the questionnaire were first screened and cleaned. The assumptions were examined before the inferential statistical analyses were conducted.

3.6.4.2.1 Mixed between-within subjects analysis of variance (Mixed ANOVA)

Data collected in the quantitative stage of Phase II were subsequently subjected to Mixed ANOVA to explore whether there were significant differences in the main and interaction effects of the independent variables. Mixed ANOVA allows researcher to combine between-subjects (two or more different groups) and within-subjects variables (one group of participants exposed to two or more conditions) in one analysis (Field, 2009; Pallant, 2013). In this research, the quantitative stage of Phase II was designed to answer two of the research questions: RQ2 and RQ3 (i.e., RQ3a and RQ3b) (see Section 1.4).

For RQ2, the between-subjects independent variables were three planning conditions: PTP, OLP and CG. The within-subjects independent variable was cognitive-task variability with simple and complex levels. The dependent variables were syntactic complexity, lexical complexity, accuracy, fluency and adequacy of students' writing.

For RQ3b, the between-subjects independent variables were planning conditions (i.e., PTP, OLP and CG) and anxiety levels (i.e., high and low). The within-subjects independent variable was cognitive-task variability (i.e., simple task and complex task). The dependent variables were syntactic complexity, lexical complexity, accuracy, fluency and adequacy of students' writing.

Assumptions of the Mixed ANOVA were first examined, including distribution of normality (checked via the skewness and kurtosis, the visual inspection of histograms, and normal Q-Q plots and box plots), and Levene's *F* test of homogeneity of variances. If one of the assumptions was not satisfied, Mixed ANOVA would not be performed, and instead, only follow-up analyses (i.e., paired-samples *t*-test or Wilcoxon Signed Ranks tests and one-way ANOVAs) were conducted.

3.6.4.2.2 Post-hoc test

If a statistically significant main effect or interaction effect was detected by Mixed ANOVA, a post-hoc test was needed to find out which of the groups were statistically significantly different from one another. In this study, after each Mixed ANOVA, further follow-up paired-samples *t*-tests for within-subjects differences and one-way ANOVA for between-subjects differences were computed to explore where the differences were located.

For both RQ2 and RQ3, there were more than three between-subjects groups. One-way ANOVA was, therefore, used in the post-hoc test, since it can be used "when you have one independent (grouping) variable with three or more levels (groups) and one dependent continuous variable" (Pallant, 2013, p. 259).

3.6.4.2.3 Cluster analysis

Cluster analysis is to group "a relatively small number of groups or cluster of objects or individuals which resemble each other and which are different in some respects from individuals in other clusters" (Everitt, Landau, Leese, & Stahl, 2011, p. 13). Cluster analysis was used in the present research to identify homogenous groups of participants who were at a similar level of writing anxiety. A two-step cluster analysis in SPSS was conducted to classify the 201 participants in the three planning groups as a low, medium or high level of anxiety, based on students' responses to the SLWAI. The SPSS showed that the cluster quality of the grouping was good. Learners whose anxiety score fell within the medium-anxiety group were not included in the analysis.

3.6.5 Ensuring the reliability of the quantitative data analysis

Three raters analysed different portions of the quantitative data in order to establish reliability estimates. One of the three raters was the researcher, who rated all parts of the data. For the other two raters, one was an English writing lecturer who has been an ESL instructor in a Chinese university for over five years. Twenty percent of the written data were given to her for analysis of accuracy. The identified errors marked by her and the ones identified by the researcher were compared. The intraclass correlation coefficient reached 96.7%. This rater also rated 20% of the pre-test based on Jacobs et al.'s (1981) rating scheme and the Cronbach's Alpha was .856.

The third rater was a doctoral student in the Faculty of Education and Social Work of the University of Auckland whose research interest is second language writing and who has been an IELTS writing instructor for several years. She was given 20% of the data for the two writing tasks. The researcher and the rater were blind to the conditions from which the essays were produced when grading. The two raters first independently rated 10 essays which were randomly selected from the data pool after we examined and discussed Jacobs et al. 's (1981) rating scheme. The inter-rater reliability calculated with intraclass correlation coefficient was determined to be inadequate (.421). We then conducted a think-aloud strategy when re-marking one of the 10 marked essays to identify differences among us.

We found two reasons that led to low reliability:

(a) The overlap between *content* and *organisation* criteria and between *vocabulary* and *language use* criteria mainly contributes to the different scoring. We, therefore, specified the criteria to avoid the overlapping mark. For example, keeping to the word limits belongs to *organisation* criteria rather than *content*. If students write too few words, they will lose marks in *organisation* part. The lexical errors, such as word choice, usage and collocations errors belong to the *vocabulary* part instead of *language use* criteria with grammatical errors. All the spelling and paragraphing errors belong to mechanics.

(b) The two raters had different emphases on the *content* and *language use* criteria. For *content*, the researcher focused on whether the key features and requirements were covered, while the other rater concentrated more on how students develop the topic by using various methods of argument and whether the statements were developed and supported well. With *language use*, the researcher focused on the microstructure, such as grammatical errors when grading, while the other rater put more emphasis on the macrostructure, such as readability and understandability of the essay.

To solve these problems, we discussed the criteria again and reached a consensus for the criteria and grading process. We decided to read each composition twice in approximately 10 minutes, when grading.

In our first reading, we assessed the readability of the essay (language and vocabulary from a macro perspective) and rated the *content* and *organisation* of the essay. In the second reading, we graded the *language*, *vocabulary* (in a micro perspective) and *mechanics*, and we reconfirmed whether the *content* and *organisation* scores were as effective as we had thought. The raters then took the remaining essays home to rate independently. As Jacobs et al. (1981) suggested, the essay would be re-scored by a third rater, when the scores given by the two raters differed by more than or equal to 9 marks. Instead of the involvement of a third independent rater, the researcher re-rated the essays with the divergence in scoring (the second round of the scoring by the researcher). The percentage of essays that was re-rated on the third occasion was only 3.98%. After resolving the differences in scores through the third rating, the intraclass correlation coefficient reached .881, with content .760 and organisation .831.

To establish intra-rater reliability, the researcher rated 402 writing tasks on two rounds with a two-month interval. Following the suggestion of Jacobs et al. (1981), for scores given by the two rounds of the researcher that differed by less than 9 marks, the average score of the two ratings was used as the operational score. When the score obtained in the two rounds deviated from each other by more than, or equal to, 9 marks, the researcher re-rated the essay on the third round. The two closest scores were then used and averaged. The percentage of essays that was re-rated on the third occasion was 5.47%. After resolving the differences in scores through the third rating, the intraclass correlation coefficient was high, .863 for the total score, .811 for the content and .793 for the organisation. The final score for content, organisation and overall score was an average of the two scores received in the two rounds.

3.7 Phase II—Qualitative Stage of the Main Study

The purpose of the qualitative stage is to provide an in-depth understanding of the quantitative data in Stage II of Phase II and to probe participants' perceptions of language writing anxiety related to different levels of cognitive-task variability and planning conditions. The qualitative stage of Phase II was guided by the following research question:

What are the relationships among students' perceived anxiety, cognitive-task variability, planning type and EFL writing production?

(c) What are students' perceptions of second language writing anxiety related to cognitive-task variability?

(d) What are students' perceptions of second language writing anxiety related to planning type?

Students were initially sorted by anxiety levels based on their responses to SLWAI from the lowest to highest score. Participants for the interview were purposively selected from the pool (students showed their willingness to participate in the interview) according to their anxiety points ranking in the top highest or lowest. Consequently, participants were 3 females in the OLP-H group, 2 females in the OLP-L group, 3 females in the PTP-H group, 2 males in the PTP-L group, 1 male and 1 female in the CG-H group and 3 females in the CG-L group.

The interviews were conducted on the day after students had completed the second writing task, as Dörnyei (2007) suggests that the interval between the task and the retrospective interview be kept as short as possible. In this way, the process of task completion was still fresh in the participant's minds. Interview questions concerned students' perceptions of L2 writing anxiety related to different levels of cognitive-task variability and different planning types. Appendix E shows the interview questions for each group.

The participants were interviewed individually in a comfortable place (an empty classroom). All the contents or questions were explained in Chinese. Each interview lasted for about 20 minutes. Students were informed that their responses were recorded and they had rights not to answer any questions they did not wish to answer and to have the recorder switched off at any time.

3.7.1 Data analysis

The audio-recordings of student interviews were transcribed verbatim by the researcher, and the transcripts were sent to the participants to check, and to amend and delete any statements. Afterwards, the qualitative data were analysed in a recursive and iterative manner with the researcher moving between the data and the research questions and the theoretical framework. The thematic analysis and supplementary content analysis were conducted in this study for the qualitative data. Drawing on a general analysis procedure which is endorsed by many researchers (Creswell, 2013; Dörnyei, 2007), the researcher, first, read and re-read the transcription of each participant to obtain a general sense of the data. Memos, reflective notes and interesting features were written down in the margins to record the key concepts and ideas emerged. This familiarization stage was done in a free and causal flowing manner without any stereotype of predetermined questions or theoretical framework. The relevant pieces were labelled, and the data were gradually segmented into meaningful chunks. During the process of creating codes, deductive and inductive coding approaches were employed. For some data, such as students' focus during the planning and writing process, a list of a priori codes were initially developed, drawing on previous studies and the theoretical framework, while new codes could emerge. For other data, such as the perceptions and sources of anxiety, and perceptions of cognitive-task variability and planning type, an inductive coding approach was used with the analytic process starting from the data itself. As the codes identified from the data grew in number, they were scanned and compared more closely to classify and cluster them to create a more general and abstract level theme.

As well as the thematic analysis, a supplementary analysis, content analysis was also adopted for the qualitative stage, which involved "the counting of instances of words, phrases, or grammatical structures that fall into specific categories" (Dörnyei, 2007, p. 245). Content analysis was used to present different proportions of students' preference and perceptions, such as the percentage of the sources of writing anxiety, and different foci during writing and planning.

3.7.2 Trustworthiness of the qualitative data analysis

Trustworthiness is a fundamental concern in qualitative research, which is a parallel concept to reliability and validity in quantitative research (Creswell, 2013). To increase the trustworthiness of the qualitative stage, participants were, first, interviewed in the language with which they felt most comfortable. Second, the interview transcripts and interpretation of the data were sent back to the participants for member checking. Third, participants, the research setting and research findings of the qualitative phase were described with sufficient details and examples from the data. Fourth, engaging in the process of data collection and analysis, the researcher's supervisor was qualified for the role as an auditor. Five, a doctoral student majoring in applied linguistics, who speaks Chinese as her mother tongue, was invited to evaluate the preliminary coding and her doubts or questions were addressed.

3.8 Ethical Considerations

This study was conducted after obtaining approval from The University of Auckland Human Participants Ethics Committee on 10 November 2017 (Reference Number 020135). Prior to the fieldwork, the researcher started contacting the Dean of the relevant faculty in the potential universities in Shandong province, China. The Participant Information Sheet (PIS) and Consent Form (CF) were sent to them with a full description of the aims of the study, the time commitment, participating requirement and any other necessary information. After gaining permission to contact teachers and students, the researcher met staff members at the two universities to explain the research, deliver PISs and CFs, and invite participation. Participants recruited in this study received the PIS and CF for their signatures to record their consent.

Before the study, participants were informed about this research and that their participation was voluntary. The participants could withdraw from the study, or withdraw any data they provided without giving any reasons up to three weeks after data collection was completed, without affecting their normal courses, grades and relationships with their teachers. They were also assured that they had the right not to answer any specific questions without being disadvantaged.

Participants were given assurance of anonymity and provided with an identification code as a serial number; they were asked to use this code rather than their names, or include any identifiable information on the questionnaires and writing tasks. The information from questionnaires, writing tasks and participants' identifiers were linked by the code number. This identifying information was kept separate from the research data and known only to the researcher. They were assured that no identifying information would be disclosed to a third party.

Regarding data management, hard copy data were safely stored in a locked cabinet at the University of Auckland. Electronic data were securely and confidentially stored on the researcher's computer, which was password-protected. All data will be shredded, and the digital information will be deleted permanently after six years. The data collected would be used only for the researcher's PhD thesis, academic publications and conference presentations, where anything related to this research, including the participating institutions and students, will be presented anonymously.

3.9 Chapter Summary

This chapter has justified the methodology and methods adopted in this research. An overview of the mixed-methods design and a summary of the results of instruments piloting and validation were presented at first. Afterwards, the information on participants, the use of instruments, the procedure of data collection of each part of the research and the analysis of both quantitative and qualitative data were carefully explained. This chapter closed with a discussion of the ethical consideration of the study. In the next chapter, findings for RQ1 and RQ2 are reported. Results for RQ3 are reported in Chapter Five.

Chapter 4 Validation of Cognitive-task Variability and Effects of Cognitive-task Variability and Planning Type

4.1 Overview

This chapter presents the findings for the validation of cognitive-task variability and for the effects of cognitive-task variability and planning type on EFL writing performance. Section 4.2 addresses RQ1: *What are the relationships between presumed cognitive-task variability and cognitive load?* The results for RQ2 are reported in Section 4.3: *What are the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing?* In each section, the background information for the participants under the different conditions is outlined, and the quantitative findings and summaries for each research question are, then, presented.

4.2 Validation of Cognitive-task Variability

Findings of Phase I of the main study were reported in this section. Quantitative data collected from dual-task experiments and TDMEQ were analysed to validate the cognitive-task variability of the two argumentative tasks. In phase I of the main study, participants were put into three groups: two dual-task groups and a single-task group. For dual-task groups, participants completed the tasks under two conditions, and each of the two groups conducted different versions of the secondary task (see more details in Section 3.5).

Before the statistical analyses, the assumption of normality was first checked via the skewness and kurtosis, the visual inspection of histograms, and normal Q-Q plots and box plots. Data are assumed to be normally distributed when the standardised skewness values are between 0 and +/- 3.0 and standardised kurtosis values between 0 and +/- 8.0 (Field, 2009; Kline, 2011). Paired-samples *t*-tests and a series of one-way repeated-measures ANOVA were computed in the following sections to examine whether the manipulation of cognitive-task variability was reasonable using self-rating scales and dual-task methodology.

4.2.1 Results for self-ratings

Table 4.1 displays the descriptive results for TDMEQ in the simple and complex task versions under the single-task condition. As Table 4.1 shows the complex task was rated, on average, to be more difficult and mental-effort consuming than the simple one.

Table 4.1 Descriptive Statistics and Paired-samples *t*-tests for Self-ratings of Perceived Task Difficulty and Mental Effort

	Simple task		Complex task		*t* (S-C)	Sig.(2-tailed)	*d*
	M	*SD*	*M*	*SD*			
ME (Item 1)	5.00	1.34	5.61	1.56	-2.386	.024**	-.418
TD (Item 3)	4.71	1.13	5.35	1.33	-2.477	.019**	-.521

Note. ME = item rated on mental effort, TD = item rated on task difficulty; S-C = the score on the simple task-the score on the complex task.
$^{**}p<.05$; $^{***}p<.001$

As the assumption of normality was met, separate paired-samples *t*-tests were computed to examine whether there were significant differences in self-ratings of task difficulty and mental effort between the simple task and the complex task.

As expected (see Table 4.1), participants perceived the complex task (M = 5.61, SD = 1.56) as significantly more mental effort required: t (30) = -2.386, p = .024, d = -.418, compared with the simple task (M = 5.00, SD = 1.34), and the effect size was medium, based on Cohen's (1988) guidelines[8]. The same was true for participants' self-ratings on task difficulty for the simple task version (M = 4.71, SD = 1.13) versus the complex version (M = 5.35, SD = 1.33), t (30) = -2.477, p = .019, d = -.521. The effect size was medium. These findings indicate that the designed-to-be complex task was more complex than the designed-to-be simple task in terms of self-perceived task difficulty

[8] The Cohen's *d* was calculated by the ES_Calculator, downloaded from http://mason.gmu.edu/~dwilsonb/downloads/ES_Calculator.xls. According to Cohen (1988), 0.2 represents a "small" effect size, 0.5 a "medium" effect size and 0.8 a "large" effect size.

and mental effort; this was found to be well aligned with the judgements of the participants (n = 5) and experts (n = 7) in the pilot study (see Section 3.4.2.1).

Pearson's correlation coefficients were also calculated, indicating a strong but not perfect correlation between perceived task difficulty and mental effort (simple task: $r = .55$, $p = .001$; complex task: $r = .61$, $p < .001$). The results suggested that the two variables were related but not isomorphic (Brünken et al., 2010; Révész et al., 2016).

In summary, the results for self-ratings confirmed the validity of the researcher's manipulations of the simple and complex tasks employed in this study, that is, the designed-to-be complex task was perceived by participants as more difficult and more mental effort consuming, compared with the simple version.

4.2.2 Results for dual-task Experiment 1

Table 4.2 summarizes the descriptive statistics for reaction time, accuracy rate and two types of error rates (Error I and Error II) on the secondary task (*di* and *du* version) under the baseline, simple, and complex task conditions. It can be seen that participants responded to the sound stimuli fastest under the baseline condition and the slowest when engaging in the complex task. The accuracy rate of the participants' reaction decreased as the task condition became more difficult. In contrast, error rates (Error I and Error II) increased with the task condition getting complex.

Table 4.2 Descriptive Statistics for Reaction Time, Accuracy Rates and Error Rates in Dual-task Experiment 1

	Baseline		Simple		Complex	
	M	*SD*	*M*	*SD*	*M*	*SD*
RT (ms)	543.91	123.12	848.13	189.61	934.77	311.20
Acc (%)	98.11	2.17	94.19	4.26	91.60	4.62
Err I (%)	1.01	1.70	3.57	3.84	3.98	3.82
Err II (%)	1.62	3.01	4.74	7.60	8.99	5.69

Note. RT = reaction time, ms = milliseconds, Acc = accuracy rates, Err I = Error I, Err II = Error II.

To explore whether participants' performance in secondary tasks differed across the baseline, simple and complex task conditions, a series of one-way repeated-measures

ANOVA was conducted. The assumptions of normality and Sphericity were first examined and found to be satisfied.

The ANOVA tests showed that each of the four analyses yielded a significant and large effect for participants' performance in the secondary task across three task conditions, reaction time: F (2, 58) = 37.892, p <.001, η_p^2 = .566; Error I: F (2, 58) = 7.436, p = .001, η_p^2 = .204; Error II: F (2, 58) = 13.807, p<.001, η_p^2 = .323; accuracy rates: F (2, 58) = 21.144, p <.001, η_p^2 = .422. These results suggest that there were significant differences in reaction time, accuracy, and error rates across baseline, simple and complex conditions. Pairwise comparisons with Bonferroni adjustment were, therefore, performed to investigate where the statistically significant differences were.

Table 4.3 Post-hoc Comparisons for Reaction Time, Accuracy, Error I, Error II across Baseline, Simple and Complex Conditions

Task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
RT	BL-ST	-304.23	40.04	.000***	-1.898
	BL-CT	-390.86	55.14	.000***	-1.563
	ST-CT	-86.63	45.04	.193	-.311
Acc	BL-ST	3.93	1.01	.002**	1.194
	BL-CT	6.51	1.05	.000***	1.861
	ST-CT	2.59	.97	.038**	.581
Err I	BL-ST	-2.56	.85	.017**	-.891
	BL-CT	-2.98	.84	.004**	-1.035
	ST-CT	-.41	.82	1.000	-.108
Err II	BL-ST	-.312	1.54	.154	-.548
	BL-CT	-7.37	1.09	.000***	-1.592
	ST-CT	-4.25	1.55	.031**	-.630

Note. RT = reaction time, Acc = accuracy rates, Err I = Error I, Err II = Error II, BL = baseline condition, ST = simple task condition, CT = complex task condition.
***p*<.05; ****p*<.001

For reaction time (see Table 4.3), participants, as expected, reacted to the secondary task significantly faster under the baseline condition (M = 543.91, SD = 123.12) than either the simple condition (M = 848.13, SD = 189.61): p <.001, d = -1.898, or the complex condition (M = 934.77, SD = 311.20): p <.001, d = -1.563; the effect sizes were large. In contrast, increases in the cognitive-task variability of the primary writing task did not have a significant impact on participants' reaction time, $p = .193$, $d = -.311$. These results indicate that the design of the dual-task methodology was sufficiently sensitive, in terms of reaction time, to distinguish the baseline and experimental conditions. Increasing cognitive-task variability of the primary tasks, however, did not statistically affect participants' secondary task reaction time.

Significant differences were found in the results of Error I between the baseline (M = 1.01, SD = 1.70) and simple conditions (M = 3.57, SD = 3.84): $p = .017$, $d = -.891$, and between baseline and complex conditions (M = 3.98, SD = 3.82) $p = .004$, $d = -1.035$. The effect sizes were large. Nevertheless, no significant differences in participants' Error I rates were found between simple and complex condition: $p = 1.000$, $d = -.108$. No significant differences for Error II were detected between the baseline and simple conditions: $p = .154$, $d = -.548$, although the effect size was medium. Nonetheless, participants made significantly more errors under the complex condition (M = 8.99, SD = 5.69) than either under baseline condition (M = 1.62, SD = 3.01): $p<.001$, $d = -1.592$ with a large effect size; or simple condition (M = 4.74, SD = 7.60): $p = .031$, $d = -.630$, with a medium to large effect size. These findings partially support the validity of cognitive-task variability manipulation. The relatively low error rates of Error I and Error II support the validity of the reaction time data. The results for error rates suggest that the complex primary task contributed to participants' more Error II in the secondary task. It appears that the baseline and experimental conditions can be sensitively distinguished by Error I and partly differentiated by Error II.

Significantly higher accuracy rates were produced on the secondary task when participants were under the simple task condition (M = 94.19, SD = 4.26), as compared with the complex task (M = 91.60, SD = 4.62), $p = .038$, $d = .581$. The effect size was medium. Participants under the baseline condition (M = 98.11, SD = 2.17) outperformed those under either the simple task condition, $p = .002$, $d = 1.194$, or the complex condition, $p<.001$, $d = 1.861$; the effect sizes were large. The results reveal that accuracy as a measure of students' cognitive load can sensitively separate the baseline and experimental conditions and predict participants' greater consumption of cognitive load when the cognitive-task variability of the primary tasks was increased.

In sum, with the exception of Error II rates, significant differences were identified between the baseline and simple conditions, and between the baseline and complex conditions, which indicates the design to measure cognitive load in dual-task Experiment 1 was generally reasonable and sensitive. Furthermore, participants obtained significantly higher accuracy rates on the secondary task when carrying out the simple primary writing task, although no evidence of reaction time was detected to support the manipulation of cognitive-task variability.

4.2.3 Results for dual-task Experiment 2

Participants' reaction time and accuracy rates were calculated, based on the design of the auditory secondary task with only *di* stimulus (shown in Table 4.4).

Table 4.4 Descriptive Statistics for Reaction Time, and Accuracy Rates in Dual-task Experiment 2

	Baseline		Simple		Complex	
	M	SD	M	SD	M	SD
RT (ms)	385.14	65.33	496.95	86.50	533.37	104.71
Acc (%)	97.88	3.14	95.47	2.76	94.04	3.10

Note. RT = reaction time, ms = milliseconds, Acc = accuracy rates.

The ANOVA results showed that the effect of task conditions (i.e., baseline, simple and complex task conditions) on participants' reaction time of the secondary task was significant, F (2, 58) = 67.660, p<.001, η_p^2 = .700, and with a large effect size. As for accuracy, a significant effect of task conditions on participants' performance was identified, F (2, 58) = 16.576, p<.001, η_p^2 = .364, and associated with a large effect size. Pairwise comparisons of estimated marginal means with Bonferroni adjusted α level were applied to investigate where the statistically significant differences were located (see Table 4.5).

Table 4.5 Post-hoc Comparisons for Reaction Time and Accuracy across Baseline, Simple and Complex Conditions

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
RT	BL-ST	-111.82	13.28	.000***	-1.424
	BL-CT	-148.24	15.15	.000***	-1.578
	ST-CT	-36.42	11.10	.008**	-.366
Acc	BL-ST	2.41	.56	.001**	.812
	BL-CT	3.84	.80	.000***	1.229
	ST-CT	1.43	.64	.103	.485

Note. RT = reaction time, Acc = accuracy rates, BL = baseline condition, ST = simple task condition, CT = complex task condition.
**$p<.05$; ***$p<.001$

The pairwise comparisons revealed that increasing cognitive-task variability of the primary writing task had a significant impact on participants' reaction time in the secondary task: Simple task (*M* = 496.95, *SD* = 86.50) versus complex task (*M* = 533.37, *SD* = 104.71), *p* = .008, *d* = -.366, with an effect size that ranged from small to medium. Participants also responded to the sound significantly faster under the baseline condition (*M* = 385.14, *SD* = 65.33), than under either the simple task condition, $p<.001$, *d* = -1.424, or the complex task condition, $p<.001$, *d* = -1.578; the effect sizes were large. These findings indicate that increasing cognitive-task variability of the primary writing task statistically increased participants' reaction time in the secondary task. Moreover, the design of this version of the secondary task was reasonable and sensitive in terms of reaction time to distinguish the baseline and experimental conditions.

For accuracy rates, increasing the cognitive-task variability of the primary task, in contrast to our expectation, did not contribute to a significant decline: Accuracy for the simple task (*M* = 95.47, *SD* = 2.76) was similar to that for the complex task (*M* = 94.04, *SD* = 3.10), *p* = .103, *d* = .485. Significant differences were, however, detected between the baseline (*M* = 97.88, *SD* = 3.14) and the simple conditions, *p* = .001, *d* = .812 and between the baseline and the complex conditions, $p<.001$, *d* = 1.229; the effect sizes were large. The results suggest that the cognitive-task variability does not statistically affect accuracy, but that baseline and experimental conditions are differentiated by accuracy.

In sum, the comparison results for the baseline-simple and baseline-complex conditions indicate that the combination of the two primary tasks and the auditory secondary task in Experiment 2 constituted a relatively sensitive measure of cognitive load. The results for reaction time of the secondary tasks offer evidence in support of the manipulation of two writing tasks; that is the designed-to-be complex task was more complex. Although the results for accuracy rates in this experiment were not supportive of the cognitive-task variability manipulation, the relatively high accuracy rates support the validity of the reaction time data.

4.2.4 Summary

Increasing cognitive-task variability yielded a significantly greater self-perceived difficulty rating and consumed participants' more mental efforts, thus confirming the findings in the pilot study.

The results of two dual-task experiments also confirm the validity of cognitive-task variability manipulation. The significant differences in reaction time, accuracy and error rates for the baseline-simple and baseline-complex conditions indicate that the dual-task design in the two experiments was reasonable and sensitive to differences in cognitive demand for the participants.

In dual-task Experiment 1, there were significantly higher accuracy rates of response under the simple primary task condition than the complex task. Participants also made significantly more errors under the complex condition than the simple one. Although the baseline-simple comparison for Error II was not significant, the validity of cognitive-task variability manipulations was supported partially. In dual-task Experiment 2, there was significantly greater reaction time under the complex primary task condition than the simple condition. Findings of the two dual-task experiments were well aligned with the intended cognitive-task variability design: Accuracy rates in Experiment 1, and reaction times in Experiment 2 predicted that participants would consume more cognitive load when completing the complex primary writing task than the simple one.

Overall, the results for self-ratings and two dual-task experiments indicate that the cognitive-task variability manipulations in this research were successful: The complex task version, as intended, yielded greater cognitive demands than the simple version.

4.3 Effects of Cognitive-task Variability and Planning Type

This section reports findings of RQ2 that examines the effects of cognitive-task variability and planning type on EFL writing. To establish whether the differences in the pre-test performance across the three planning groups were statistically significant, a one-way ANOVA was performed. The means and standard deviations for pre-tests in each planning group are shown in Table 4.6; no statistically significant group differences in pre-writing test among the three groups were found, F (2, 198) = .554, p = .575.

Table 4.6 Group Means and Standard Deviations for Pre-writing Test

	N	*M*	*SD*
OLP	68	61.44	3.85
PTP	68	60.72	3.66
CG	65	61.23	4.74

There were six conditions in total: Simple-comparison (S-CG), complex-comparison (C-CG), simple-online planning (S-OLP), complex-online planning (C-OLP), simple-pre-task planning (S-PTP), and complex-pre-task planning (C-PTP). Participants' writing performance was analysed in terms of syntactic and lexical complexity, accuracy, fluency, content, organisation and overall text quality.

Before the statistical analyses, the assumption of the normal distribution for each condition was first checked via the skewness and kurtosis, the visual inspection of histograms, and normal Q-Q plots and box plots. Furthermore, the assumption of homogeneity of variances was evaluated based on Levene's F test. When all the assumptions were satisfied, two-way mixed ANOVAs were conducted to explore the main effects, and interaction effects of cognitive-task variability (i.e., simple vs. complex) and planning type (i.e., PTP, OLP and CG) on EFL writing production. Further follow-up paired-samples t-test for within-subjects differences (i.e., differences between simple and complex tasks) and one-way ANOVA for between-subjects differences (i.e., differences across three planning types) were computed to explore where the differences were located. If either the assumption of normality, or the assumption of homogeneity of variances was violated, mixed-ANOVA would not be performed, and only follow-up analyses were executed for these variables. In the following sections, the results of the diagnostic tests for each group were reported prior to presenting the main findings.

4.3.1 Effects on syntactic complexity

Findings of the effects of cognitive-task variability and planning type on students' syntactic complexity are reported in this section. EFL syntactic complexity, a multidimensional conceptualisation (Norris & Ortega, 2009), was measured in four sub-constructs: general complexity (MLT), complexity by subordination (C/T, and DC/C), complexity via coordination (T/S) and complexity via phrasal elaboration (MLC, CP/C, and CN/C). Table 4.7 displays the descriptive results for the seven syntactic complexity measures in the simple and complex tasks across the CG, OLP and PTP groups.

Table 4.7 Descriptive Statistics for Syntactic Complexity across Three Planning Groups in Two Tasks

Syntactic complexity	Groups	*N*	Simple task		Complex task	
			M	*SD*	*M*	*SD*
MLT	CG	65	12.32	2.44	11.95	3.00
	OLP	68	12.42	2.47	11.82	2.52
	PTP	68	11.93	1.91	11.84	2.68
C/T	CG	65	1.47	.27	1.46	.32
	OLP	68	1.52	.26	1.41	.23
	PTP	68	1.43	.19	1.41	.21
DC/C	CG	65	.30	.11	.28	.12
	OLP	68	.33	.10	.28	.10
	PTP	68	.29	.09	.27	.10
T/S	CG	65	1.12	.09	1.13	.12
	OLP	68	1.10	.09	1.11	.10
	PTP	68	1.12	.13	1.11	.10
MLC	CG	65	8.40	1.11	8.21	1.00
	OLP	68	8.17	.88	8.38	1.14
	PTP	68	8.39	1.25	8.39	1.22
CP/C	CG	65	.32	.13	.36	.16
	OLP	68	.31	.11	.36	.13
	PTP	68	.33	.12	.37	.12
CN/C	CG	65	.87	.21	.83	.22
	OLP	68	.85	.19	.84	.23
	PTP	68	.87	.22	.83	.24

Note. PTP=pre-task planning group, OLP=online planning group, CG=comparison group, MLT=mean length of T-unit, C/T=clauses per T-unit, DC/C=dependent clauses per clause, T/S=T-units per sentence, MLC= mean length of clause, CP/C=coordinate phrases per clause, CN/C=complex nominals per clause.

4.3.1.1 Effects on MLT

Figure 4.1 depicts the mean scores of MLT in the simple and complex tasks across the CG, OLP and PTP groups. As the graph shows, the scores of all three groups declined when cognitive-task variability was increased, with the OLP group demonstrating the biggest decrease. In the simple task, students under the OLP condition produced the most MLT, followed by those under the CG and PTP conditions. In the complex task, MLT of the CG and the PTP group outnumbered that of the OLP group.

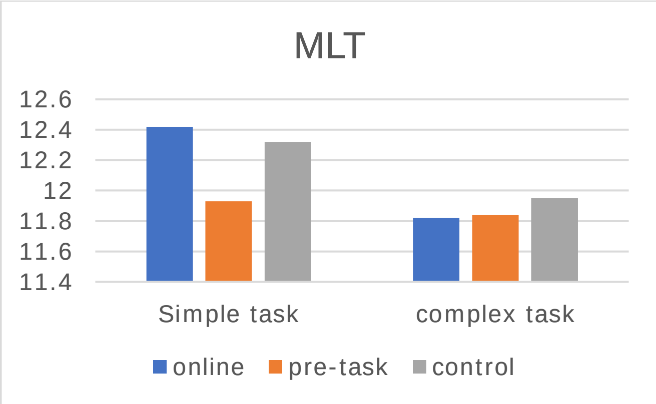

Figure 4.1 Mean MLT for simple and complex tasks across three planning conditions

To examine the effects of cognitive-task variability and planning type on MLT as a measure of students' overall syntactic complexity, the assumption of normality was checked and found to be met. Levene's *F* test for MLT in the simple task, *F* (2, 198) = 2.250, *p* = .108, and for MLT in the complex task, *F* (2, 198) = .014, *p* = .986, was not significant, indicating that the variance of the dependent variable, in either the simple or the complex task, was equal across the groups. Therefore, a mixed ANOVA was applied.

Table 4.8 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on MLT

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	4.996	.027**	.025
group	2	198	.260	.772	.003
task*group	2	396	.866	.422	.009

Note. Task = cognitive-task variability, group = planning type.
***p*<.05; ****p*<.001

The mixed ANOVA results are shown in Table 4.8[9]. There was a significant main effect of cognitive-task variability, *F* (1, 396) = 4.996, *p* = .027, η_p^2 = .025, but neither main effect for planning type *F* (2, 198) = .260, *p* = .772, η_p^2 = .003, nor interaction between

[9] The alpha value was set at .05 for the report of results of this study, except that for the Box's test, which was set at .001.

cognitive-task variability and planning type, *F*(2, 396) = .866, *p* = .422, η_p^2 = .009, was significant. These findings indicate that increasing cognitive-task variability (if the planning type variable was ignored) decreased the number of MLT students made, and varying planning type did not significantly affect their MLT production (if the cognitive-task variability was ignored). Students' MLT production in the simple versus the complex versions was not dependent on planning type.

To explore the difference between participants' MLT production on the simple task versus complex task versions respectively under the CG, OLP, and PTP conditions, three paired-samples *t*-tests were performed as reported in Table 4.9.

Table 4.9 Paired-samples *t*-tests for MLT across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.37	.24	1.538	.129	.132
OLP (S-C)	68	.60	.31	1.909	.061	.238
PTP (S-C)	68	.09	.26	.353	.725	.038

Note. MLT = mean length of T-unit, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.

Paired-samples *t*-tests revealed that in the CG, no significant difference was found in MLT scores between the simple (*M* = 12.32, *SD* = 2.44) and complex tasks (*M* = 11.95, *SD* = 3.00), *t* (64) = 1.538, *p* = .129, *d* = .132. Likewise, the PTP group produced MLT similarly under the simple (*M* = 11.93, *SD* = 1.91) and the complex task (*M* = 11.84, *SD* = 2.68) conditions, *t* (67) = .353, *p* = .725, *d* = .038. The OLP group also made similar MLT scores between the complex task (*M* = 11.82, *SD* = 2.52) and the simple task (*M* = 12.42, *SD* = 2.47), *t* (67) = 1.909, *p* = .061, *d* = .238. These findings indicate that increasing cognitive-task variability did not lead to a significant decline in students' overall syntactic complexity under either PTP, OLP or CG condition.

Two separate one-way ANOVA tests were conducted to examine whether significant between-subjects differences (i.e. differences of MLT among three planning groups for each task) existed. The ANOVA results showed that in the simple task, participants' MLT did not differ significantly among the CG, OLP and PTP groups, *F* (2, 198) = .869, *p* = .421. The same was true for the students' MLT production across the online planning, pre-task planning and comparison groups, *F* (2, 198) = .039, *p* = .961 with the complex task. The results indicate different planning types did not significantly affect students' MLT either in the simple or in the complex version of the task.

4.3.1.2 Effects on C/T

In this section, the results for the effects of cognitive-task variability and planning type on C/T as a measure of subordination of clause were reported. Figure 4.2 depicts the mean scores of C/T. As the graph shows, both the OLP and PTP groups experienced a decrease of C/T, when the cognitive-task variability was increased. Changes of C/T scores in the comparison group were negligible.

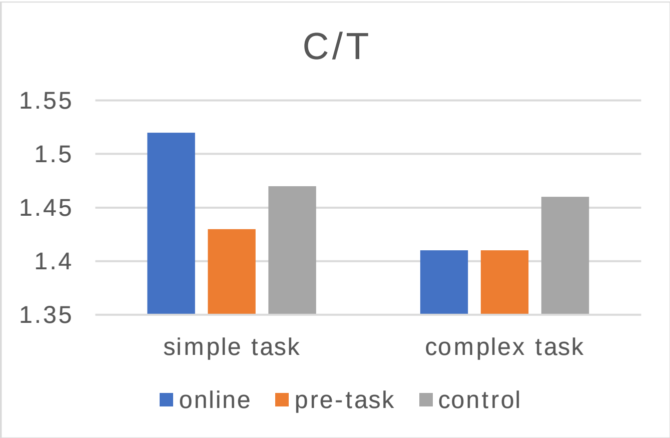

Figure 4.2 Mean C/T for simple and complex tasks across three planning conditions

The assumption of normality was met. Levene's *F* tests in both the simple task, *F* (2, 198) = 1.854, *p* = .159, and the complex task, *F* (2, 198) = 1.769, *p* = .173, showed that the homogeneity of variances assumption was not violated. Hence, a mixed ANOVA was conducted.

Table 4.10 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on C/T

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	7.517	.007**	.037
group	2	198	1.074	.344	.011
task*group	2	396	2.589	.078	.025

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

As displayed in Table 4.10, the mixed ANOVA yielded a significant main effect for cognitive-task variability, F (1, 396) = 7.517, p = .007, η_p^2 = .037. Neither main effect for planning type, F (2, 198) = 1.074, p = .344, η_p^2 = .011, nor interaction between cognitive-task variability and planning type, F (2, 396) = 2.589, p = .078, η_p^2 = .025, however, reached statistical significance. These results indicate that providing different planning types did not significantly influence students' C/T scores (without considering the effects of cognitive-task variability). Additionally, students' C/T performance was not interactively affected by cognitive-task variability and planning type. Increasing cognitive-task variability (without considering the effects of planning type), however, negatively affected participants' C/T production.

Three separate paired-samples *t*-tests were computed to evaluate the differences in the students' syntactic complexity in terms of C/T for the simple task versus the complex task versions respectively under the CG, OLP, and PTP conditions.

Table 4.11 Paired-samples *t*-tests for C/T across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.02	.03	.534	.595	.062
OLP (S-C)	68	.11	.03	3.252	.002**	.458
PTP (S-C)	68	.02	.03	.858	.394	.119

Note. C/T = clauses per T-unit, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

Table 4.11 revealed that in the CG, participants' C/T did not significantly differ in the simple task (*M* = 1.47, *SD* = .27), compared to those in the complex task (*M* = 1.46, *SD* = .32), *t* (64) = .534, *p* = .595, *d* = .062. The same pattern was found in the PTP group, where students produced C/T similarly under the simple (*M* = 1.43, *SD* = .19) and complex task conditions (*M* = 1.41, *SD* = .21), *t* (67) = .858, *p* = .394, *d* = .119. Online planners, however, used significantly more C/T in the simple task (*M* = 1.52, *SD* = .26) than in the complex task (*M* = 1.41, *SD* = .23), *t* (67) = 3.252, *p* = .002, *d* = .458, with the effect size close to medium. These findings indicate that increasing cognitive-task variability had no significant influence on the C/T scores for students who were in the comparison group and pre-task planning group, but had detrimental effects on online planners' C/T.

Two one-way ANOVA tests were performed to explore the differences in C/T production over three planning groups in both the simple and complex tasks. The ANOVA results showed that for the simple task, the between-group differences among the three planning types were not significant, *F* (2, 198) = 2.354, *p* = .098. Likewise, the three planning groups did not significantly differ in the complex task, *F* (2, 198) = .747 *p* = .475. The results suggest planning type did not significantly affect students' C/T production, when the task version was either simple or complex.

4.3.1.3 Effects on DC/C

The effects of cognitive-task variability and planning type on DC/C, a measure of subordination of dependent clause, were examined and are presented in this section. Figure 4.3 depicts the mean scores of DC/C in the simple and complex tasks over the CG, and the OLP and PTP groups. As the graph shows, a decreasing trend occurred across the three planning groups when the cognitive-task variability was increased. The PTP group produced the least DC/C in both simple and complex task versions. The OLP group made the most DC/C in the simple task, while with the complex task, the CG and OLP groups produced the same amount of DC/C.

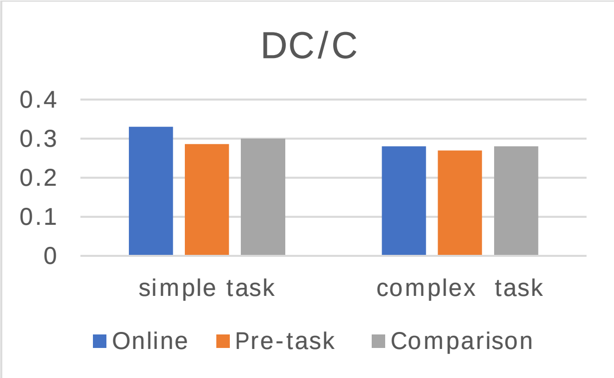

Figure 4.3 Mean DC/C for simple and complex tasks across three planning conditions

A mixed ANOVA was computed, since the assumption of normality was met and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* test for the simple task, F (2, 198) = .858, p = .426, and for the complex task, F (2, 198) = 1.357, p = .260).

Table 4.12 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on DC/C

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	11.446	.001**	.055
group	2	198	1.580	.208	.016
task*group	2	396	1.402	.249	.014

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

As shown in Table 4.12, the main effect for cognitive-task variability, F (1, 396) = 11.446, p= .001, η_p^2 = .055, but not for planning type, F (2, 198) = 1.580, p = .208, η_p^2 = .016, was found to be significant in DC/C. Interaction between cognitive-task variability and planning type in students' DC/C production, F (2, 396) = 1.402, p = .249, η_p^2 = .014, did not reach significance. These results reveal that students' DC/C production was neither significantly influenced by providing different planning types (without considering the effects of cognitive-task variability), nor affected by the interaction between cognitive-task variability and planning type. Increasing cognitive-task variability (without considering the effects of planning type), however, negatively affected students' syntactic complexity in terms of DC/C.

Table 4.13 Paired-samples *t*-tests for DC/C across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.02	.01	1.391	.169	.161
OLP (S-C)	68	.04	.01	3.228	.002**	.446
PTP (S-C)	68	.02	.01	1.197	.235	.171

Note. DC/C = dependent clauses per clause, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
$^{**}p<.05$; $^{***}p<.001$

To test the within-subjects differences in DC/C production over the OLP, PTP and CG conditions, three paired-samples *t*-tests were conducted. As Table 4.13 displays, neither the PTP group, t (67) = 1.197, p = .235, d = .171, nor the CG, t (64) = 1.391, p = .169, d = .161 produced significant differences in DC/C between the simple task (PTP: M = .29, SD = .09; CG: M = .30, SD = .11) and the complex task (PTP: M = .27, SD = .10; CG: M = .28, SD = .12). By contrast, for the OLP group, students' DC/C performance for the simple task (M = .33, SD = .10) was significantly greater than for the complex task (M = .28, SD = .10), t (67) = 3.228, p = .002, d = .446, with the effect size close to medium. The results suggest that while neither the PTP group nor the CG produced statistically more DC/C when cognitive-task variability was increased, online planning reduced participants' performance in DC/C as a function of increasing cognitive-task variability.

To explore the between-subjects differences of DC/C across the three planning groups in each task with different cognitive-task variability levels, one-way ANOVAs were conducted. According to the ANOVA results, no significant differences were found among the three planning groups when students in the complex task condition, F (2, 196) = .269, p = .764. Whereas, the OLP group (M = .33, SD = .10) almost significantly outperformed the PTP group (M = .29, SD = .09) in DC/C when students completed the simple task, p=.053, d=.423. The effect size arranged close to medium. The results indicate that students under the OLP condition had advantages over those under the PTP condition in terms of DC/C performance, when completing the simple task.

4.3.1.4 Effects on T/S

Figure 4.4 illustrates participants' T/S performance in the simple and complex tasks across CG, OLP and PTP groups. As the graph displays, the three groups differed in T/S with the tasks varying. When cognitive-task variability was increased, T/S output in the OLP group and the CG was improved, while the PTP group declined in T/S. In the simple task, the OLP group got the lowest T/S scores, and in the complex task, the CG produced the most T/S.

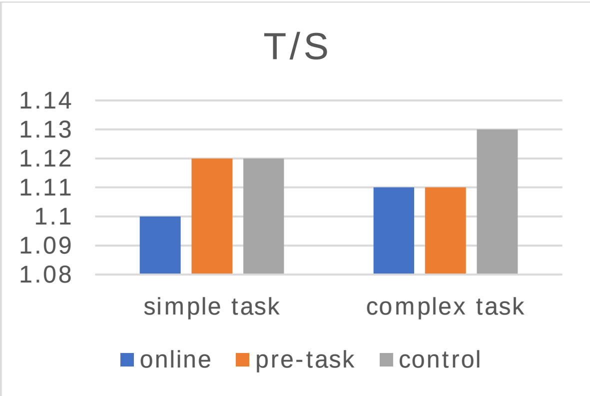

Figure 4.4 Mean T/S for simple and complex tasks across three planning conditions

Mixed ANOVA analysis was not performed, as the assumption of normal distribution in the S-PTP group for T/S was not met[10]. Other statistical analyses were employed to test the effects of cognitive-task variability and planning type on T/S.

To check the within-subjects differences in students' T/S, two paired-samples *t*-tests were conducted for the OLP group and the CG, since the assumption of normality was met. Data for the PTP group was submitted to non-parametric tests, Wilcoxon Signed Ranks tests, as the assumption of normal distribution was violated.

[10] The PTP group's T/S score in the simple task: Skewness=2.409, and kurtosis=9.531.

Table 4.14 Paired-samples *t*-tests for T/S under OLP and CG Conditions

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	-.01	.02	-.638	.526	-.09
OLP (S-C)	68	-.01	.01	-.528	.599	-.08

Note. T/S = T-units per sentence, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group.

Paired-samples *t*-tests (Table 4.14) revealed no significant differences of T/S between the simple task (*M* = 1.12, *SD* = .09) and the complex task (*M* = 1.13, *SD* = .12), *t* (64) = -.638, *p* = .526, *d* = -.09 in the CG. Likewise, the OLP group produced similar T/S under the simple (*M* = 1.10, *SD* = .09), and the complex task conditions (*M* = 1.11, *SD* = .10), *t* (67) = -.528, *p* = .599, *d* = -.08. The results suggest that increasing cognitive-task variability did not lead to significant improvements in students' T/S under the CG and OLP conditions.

Table 4.15 Wilcoxon Signed Ranks for T/S under PTP Condition

	Rank	*N*	*z*	AsympSig.(2-tailed)	*r*
S vs. C	Negative Positive	68	-.836	.403	.104[11]
	Ties	4			

Note. T/S = T-units per sentence, S = simple task, C = complex tasks.

Wilcoxon signed ranks (see Table 4.15) did not yield significant differences in participants' T/S under the PTP condition between the simple task (*M* = 1.12, *SD* = .13) and the complex task (*M* = 1.11, *SD* = .10), *z* = -.836, *p* = .403, *r* = .104. The results indicate that rising cognitive-task variability levels did not contribute to significant enhancements in students' T/S under the PTP condition.

To identify any between-subjects differences, a one-way ANOVA for the complex task and a Kruskal-Wallis *H* test for the simple task were conducted. The ANOVA results showed no significant differences in T/S production among the three planning groups when students were writing the complex task, *F* (2, 196) = .867, *p* =.422. Likewise, according to the Kruskal-Wallis *H* test, students under the OLP (*M* = 1.10, *SD* =.09),

[11] Calculated by the formula: $r = |Z|/\sqrt{N}$.

PTP (M = 1.12, SD =.13), and CG conditions (M = 1.12, SD =.09) produced T/S similarly in their simple task, $H(2)$ = 1.815, p = .404. The results reveal that providing different planning types did not enhance students' production of T/S significantly in either the simple task or the complex task.

4.3.1.5 Effects on MLC

This section reports the effects of cognitive-task variability and planning type on students' syntactic complexity via phrasal elaboration (MLC). As the Figure 4.5 illustrates, a decline occurred in the CG scores, meanwhile the OLP group showed an upward trend of MLC production when cognitive-task variability was increased by involving more elements and more reasoning demands. In the simple task, the PTP group and the CG achieved MLC to a similar level, while the OLP group produced the least MLC. With the complex task, the OLP group outperformed the CG, and produced similar MLC to the PTP group.

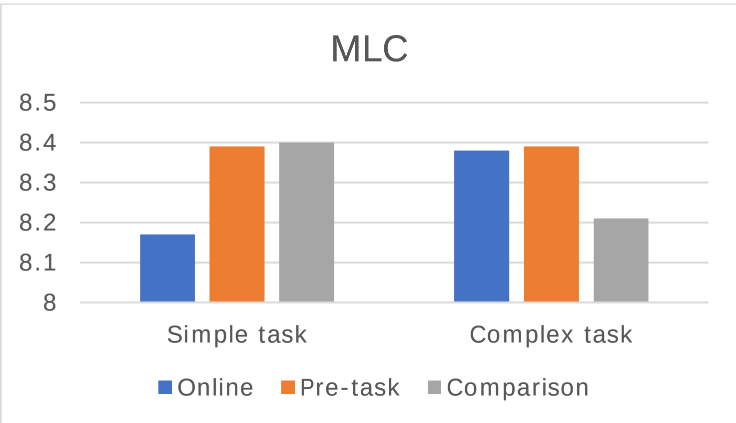

Figure 4.5 Mean MLC for simple and complex tasks across three planning conditions

The assumption of normality was first checked and found to be satisfactory. Levene's F tests for MLC in the simple task, F (2, 198) = 2.201, p = .113, and for MLC in the complex task, F (2, 198) = .436, p = .647, were not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA, therefore, was applied.

Table 4.16 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on MLC

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	.013	.910	.000
group	2	198	.297	.743	.003
task*group	2	396	2.080	.128	.021

Note. Task = cognitive-task variability, group = planning type, η_p^2 = partial η^2.
**p<.05; ***p<.001

The mixed ANOVA results for MLC are shown in Table 4.16. The main effect for neither cognitive-task variability, F (1, 396) = .013, p = .910, η_p^2 = .000, nor for planning type, F (2, 198) = .297, p = .743, η_p^2 = .003 was found for MLC production. Furthermore, no significant interaction effect on MLC, F (2, 396) = 2.080, p = .128, η_p^2 = .021 was identified. These findings indicate that increasing cognitive-task variability (if the planning type variable was ignored) did not have a significant effect on MLC; nor was participants' MLC production affected significantly by varying planning types (if the cognitive-task variability was ignored). Additionally, the students' level of MLC output in the simple versus the complex versions was not dependent on the planning types.

To explore the differences in students' MLC production between the simple and the complex task versions respectively under the OLP, PTP, CG conditions, three paired-samples t-tests were performed. As Table 4.17 displays, the differences in MLC on the simple (M = 8.40, SD = 1.11) versus the complex versions (M = 8.21, SD = 1.00) under the CG condition, t (64) = 1.518, p = .134, d = .176 were not statistically significant. The same trend was found under the PTP condition, t (67) = -.017, p = .987, d = -.002, where similar MLC scores were achieved with the simple task (M = 8.39, SD = 1.25) and the complex task (M = 8.39, SD = 1.22), and under the OLP condition, t (67) = -1.535, p = .130, d = -.206, the MLC scores did not vary markedly between the simple task (M = 8.17, SD = .88) and the complex task (M = 8.38, SD = 1.14).

Table 4.17 Paired-samples *t*-tests for MLC across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.19	.12	1.518	.134	.176
OLP (S-C)	68	-.21	.14	-1.535	.130	-.206
PTP (S-C)	68	-.003	.15	-.017	.987	-.002

Note. MLC = mean length of clause, S-C = the score on the simple task -the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.

Two separate one-way ANOVA tests were conducted to explore whether significant between-subjects differences (i.e. differences of MLC among three planning groups for each task) existed. The ANOVA results presented indicate that with the simple task, participants' MLC did not differ significantly across the OLP group (M = 8.17, SD = .88), PTP group (M = 8.39, SD = 1.25) and comparison group (M = 8.40, SD = 1.11), F (2,198) = .984, p = .376. The same was true for students' MLC production between the OLP group (M = 8.38, SD = 1.14), PTP group (M = 8.39, SD = 1.22) and comparison groups (M = 8.21, SD = 1.00) with the complex task, F (2,198) = .546, p = .580. The results suggest different planning types did not affect students' MLC either with the simple or the complex task versions.

4.3.1.6 Effects on CP/C

Figure 4.6 illustrates the trend of the mean scores across the three planning groups for CP/C. According to the graph, the three groups experienced a growing trend, when cognitive-task variability was increased. In both simple and complex tasks, the PTP group had the greatest CP/C production.

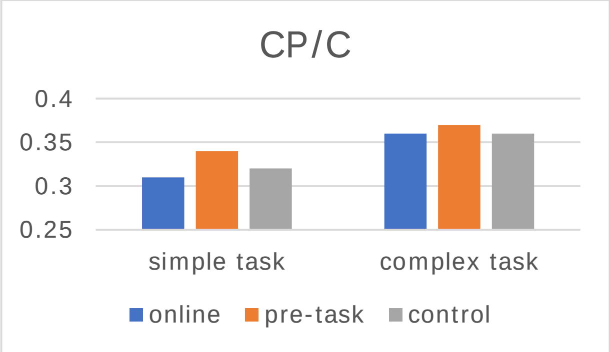

Figure 4.6 Mean CP/C for simple and complex tasks across three planning conditions

The assumption of normal distribution was checked and was satisfactory. Levene's *F* test for the simple task, F (2, 198) = 1.507, p = .224, and for the complex task, F (2, 198) = .997, p = .371, showed that the assumption of homogeneity of variances was not violated for either the simple or complex task. Hence, a mixed ANOVA was conducted to explore the effects of cognitive-task variability and planning type on students' CP/C, one of the phrase-level measures of syntactic complexity.

Table 4.18 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on CP/C

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	13.618	.000***	.064
group	2	198	.133	.876	.001
task*group	2	396	.217	.805	.002

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

As Table 4.18 shows, the mixed ANOVA yielded a significant main effect for cognitive-task variability, F (1, 396) = 13.618, p< .001, η_p^2 = .064. No statistical differences, however, were found for either the main effect for planning type F (2, 198) = .133, p = .876, η_p^2 = .001, or the interaction between cognitive-task variability and planning type F (2, 396) = .217, p = .805, η_p^2 = .002 in terms of participants' CP/C production. These results indicate that increasing cognitive-task variability had a beneficial effect on students' CP/C production (if the planning type was ignored), while the provision of different planning types did not significantly affect participants' CP/C scores (if the cognitive-task variability was ignored). Additionally, participants' CP/C performance was not influenced by the interaction of cognitive-task variability and planning type.

Three separate paired-samples *t*-tests were conducted to examine the location of the within-subjects differences in the students' CP/C output under the OLP, PTP and CG conditions respectively.

Table 4.19 Paired-samples *t*-tests for CP/C across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	-.04	.02	-1.733	.088	-.248
OLP (S-C)	68	-.05	.02	-2.919	.005**	-.427
PTP (S-C)	68	-.04	.02	-2.034	.046**	-.307

Note. CP/C = coordinate phrases per clause, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

The CG students' CP/C did not differ significantly in the simple task (M = .32, SD = .13), compared to those in the complex task (M = .36, SD = .16), t (64) = -1.733, p = .088, d = -.248 (see Table 4.19). The OLP group, however, used significantly more CP/C in the complex task (M = .36, SD = .13) than in the simple one (M = .31, SD = .11), t (67) = -2.919, p = .005, d = -.427; the effect size was close to medium. Likewise, a significant improvement occurred in the PTP group, t (67) = -2.034, p = .046, d = -.307, when the cognitive-task variability was increased from simple (M = .33, SD = .12) to complex (M = .37, SD = .12), and the effect size was small. These findings suggest that increasing cognitive-task variability had positive effects on students' CP/C production under the OLP and PTP conditions, but not under the CG condition.

Two one-way ANOVA tests were performed to explore the differences of CP/C production over three planning groups in both the simple and complex tasks. The ANOVA results identified that the between-group differences among the three planning types were not statistically significant for either simple task, F (2,198) = .277, p = .759, or complex task, F (2,198) =.032, p =.969. These findings indicate that varying the planning types affects students' CP/C output in neither the simple task, nor the complex task.

4.3.1.7 *Effects on CN/C*

Figure 4.7 illustrates the changes of CN/C between the simple and complex tasks across the CG, OLP and PTP groups. The graph presents that all the three planning groups saw a downward trend, when the cognitive-task variability was increased.

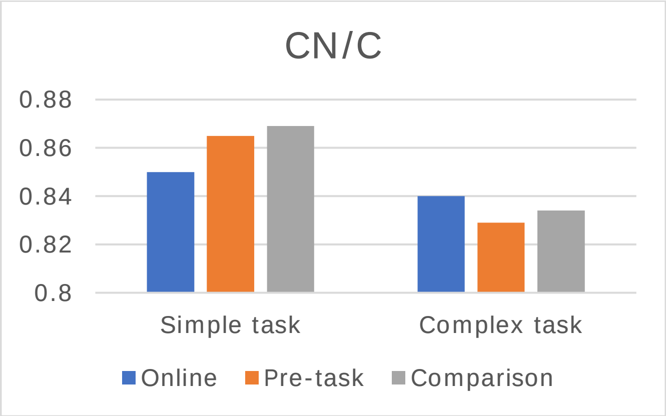

Figure 4.7 Mean CN/C for simple and complex tasks across three planning conditions

A mixed ANOVA was conducted to explore how cognitive-task variability and planning type affect CN/C, since the assumption of normality was met and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* tests in the simple task, F (2, 198) = .698, p = .499, and in the complex task, F (2, 198) = .175, p = .840.

Table 4.20 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on CN/C

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	2.898	.090	.014
group	2	198	.054	.947	.001
task*group	2	396	.254	.776	.003

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

The mixed ANOVA results (Table 4.20) showed no interaction effects between cognitive-task variability and planning type in participants' CN/C output, F (2, 396) = .254, p = .776, η_p^2 = .003. Likewise, a significant main effect for neither cognitive-task variability, F (1, 396) = 2.898, p = .090, η_p^2 = .014, nor planning type F (2, 198)

= .054, p = .947, η_p^2 = .001, in CN/C was found. These results suggest that participants' production of CN/C was not significantly influenced by either varying the planning types (without considering the effects of cognitive-task variability), or increasing cognitive-task variability (without considering the effects of planning type). Participants' CN/C production in the simple versus the complex versions was not dependent on planning types.

Table 4.21 Paired-samples *t*-tests for CN/C across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.04	.03	1.084	.282	.163
OLP (S-C)	68	.01	.03	.427	.671	.054
PTP (S-C)	68	.04	.03	1.452	.151	.157

Note. CN/C = complex nominals per clause, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

To further test the within-subjects differences in students' CN/C production respectively under the OLP, PTP and CG conditions, three separate paired-samples *t*-tests were conducted. As Table 4.21 illustrates, no significant differences in CN/C output were found between students' simple task performance (OLP: M = .85, SD = .19; PTP: M = .87, SD = .22; CG: M = .87, SD = .21) and complex task performance (OLP: M = .84, SD = .23; PTP: M = .83, SD = .24; CG: M = .83, SD = .22) regardless of planning types: OLP, t (67) = .427, p = .671, d = .054; PTP, t (67) = 1.452, p = .151, d = .157; CG, t (64) = 1.084, p = .282, d = .163. These findings reveal that increasing cognitive-task variability did not lead to any significant improvement in CN/C.

Two follow-up one-way ANOVAs were computed to examine the between-group contrasts in CN/C respectively under the simple and complex task conditions. The ANOVA results identified that students' CN/C production under the OLP, PTP and CG conditions did not differ significantly, F (2,198) = .228, p = .797, with simple task version. The complex task was similar, with no significant differences found among the CG, OLP and PTP groups, F (2,198) = .014, p = .986. These results suggest that the provision of different planning types did not contribute to significant changes in participants' CN/C output.

4.3.1.8 Summary

In this section, the effects of cognitive-task variability and planning type on syntactic complexity (seven measures) are summarized (see Table 4.22).

The within-subjects differences related to the impact of cognitive-task variability can be summarised as following:

- For the OLP group, increasing cognitive-task variability led to a significant reduction in participants’ C/T and DC/C, but marked increase in CP/C, with no significant influence on other measures of writing syntactic complexity.
- For the PTP group, the production of CP/C in the complex task substantially outnumbered those in the simple task; there were no obvious effects on other syntactic complexity measures.
- For the CG, there were no significant differences in the seven syntactic measures of complexity between the simple and complex task versions.

The between-subjects differences, as an outcome of planning type, are summarised in the following aspects. Under the simple task condition, the OLP group outperformed the PTP group in terms of students’ DC/C production, while no significant differences in syntactic complexity were found between the OLP and PTP groups under the complex task condition. None of the planning types enhanced students’ MLT, C/T, T/S, MLC, CP/C and CN/C production under either the simple or the complex task conditions.

With the simultaneous impacts of cognitive-task variability and planning type, there were no interaction effects on the seven syntactic complexity measures.

Table 4.22 Summary of Effects of Cognitive-task Variability and Planning Type on Syntactic Complexity

Syntactic complexity measures	Within-subjects differences (S vs. C)			Between-subjects differences (OLP vs. PTP vs. CG)		Interaction effects
	OLP	PTP	CG	simple task	complex task	
MLT	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.
C/T	S>C	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.
DC/C	S>C	no Sig.	no Sig.	OLP>PTP(almost)	no Sig.	no Sig.
T/S	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.
MLC	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.
CP/C	S<C	S<C	no Sig.	no Sig.	no Sig.	no Sig.
CN/C	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.	no Sig.

Note. S = Simple task, C = Complex task, OLP = online planning group, PTP = pre-task planning group, CG = comparison group, no Sig. = no significant difference.

4.3.2 Effects on lexical complexity

Lexical complexity was measured by lexical density (LD) and WT/√2W as a measure of lexical diversity. The findings of the effects of cognitive-task variability and planning type on students' lexical complexity are reported in this section. Table 4.23 displays the descriptive results for the lexical complexity in the simple and complex tasks across the CG, OLP and PTP groups.

Table 4.23 Descriptive Statistics for Lexical Complexity across Three Planning Groups in Two Tasks

Lexical complexity	Groups	*N*	Simple task		Complex task	
			M	*SD*	*M*	*SD*
LD	CG	65	.517[12]	.03	.506	.03
	OLP	68	.509	.03	.512	.03
	PTP	68	.509	.03	.510	.03
WT/√2W	CG	65	5.51	.51	5.17	.48
	OLP	68	5.56	.52	5.40	.53
	PTP	68	5.24	.46	5.21	.48

Note. PTP=pre-task planning group, OLP=online planning group, CG=comparison group, LD=lexical density, WT/√2W (a measure of lexical diversity) = the total number of different word types divided by the square root of two times the total number of words.

4.3.2.1 Effects on LD

Figure 4.8 reflects the mean scores of LD. As the graph illustrates, the trends of the three groups were different, although the means of LD in the three planning groups were very similar. Specifically, there were few changes in PTP students' LD production between the simple and complex tasks. The OLP group experienced a slight rise when the cognitive-task variability was increased. LD in the CG decreased from simple task (where students produced the most LD among the three groups) to complex task (where students produced the least LD among the three groups).

[12] Mean here was rounded to 3 decimal places, since the numbers are too close.

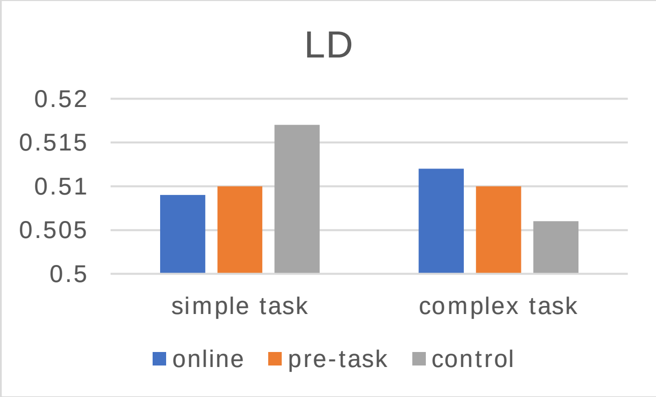

Figure 4.8 Mean LD for simple and complex tasks across three planning conditions

The data on LD was analysed with the mixed ANOVA test, since the assumption of normality was tenable and the assumption of homogeneity of variances was not violated for either the simple or complex task (Levene's F tests for LD in the simple task, F (2, 198) = .444, p = .642, and for LD in the complex task, F (2, 198) = 1.280, p = .280).

Table 4.24 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on LD

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	.933	.335	.005
group	2	198	.107	.898	.001
task*group	2	396	2.851	.060	.028

Note. Task = cognitive-task variability, group = planning type, η_p^2 = partial η^2.
**p<.05; ***p<.001

The mixed ANOVA results are presented in Table 4.24. Significant main effects were found neither for cognitive-task variability, F (1, 396) = .933, p = .335, η_p^2 = .005, nor for planning type F (2, 198) = .107, p = .898, η_p^2 = .001. Interaction between cognitive-task variability and planning type, F (2, 396) = 2.851, p = .060, η_p^2 = .028 was not significant. These findings indicate that neither increasing cognitive-task variability (if the planning type variable was ignored), nor providing planning (if the cognitive-task variability was ignored) affected students' production of LD significantly. Similarly, the LD production under simple and complex tasks did not significantly differ over the CG, OLP and PTP groups.

To locate the within-subjects differences, each level of cognitive-task variability (i.e., simple task and complex task) across the CG, OLP and PTP groups were compared. The findings of three paired-samples *t*-tests were reported in Table 4.25.

Table 4.25 Paired-samples *t*-tests for LD across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.010	.004	2.527	.014**	.335
OLP (S-C)	68	.003	.004	-.740	.462	-.097
PTP (S-C)	68	-.000[13]	.004	-.101	.919	-.014

Note. LD = lexical density, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
***p*<.05; ****p*<.001

Paired-samples *t*-tests revealed that in the CG, LD scores for students' completing the simple task (*M* = .517, *SD* = .03) were greater than for those completing the complex task (*M* = .506, *SD* = .03), *t* (64) = 2.527, *p* = .014, *d* = .335, with a small to medium effect size. However, there was no significant difference in LD between the simple (OLP: *M* = .509, *SD* = .03; PTP: *M* = .509, *SD* = .03) and complex tasks (OLP: *M* = .512, *SD* = .03; PTP: *M* = .510, *SD* = .03) under the OLP, *t* (67) = -.740, *p* = .462, *d* = -.097, and PTP conditions, *t* (67) = -.101, *p* = .919, *d* = -.014. The findings indicate that increasing the level of cognitive-task variability had a statistically negative effect on students' LD under the comparison group, but no significant impact under the OLP and PTP conditions.

Two separate one-way ANOVA tests were conducted to examine whether significant between-subjects differences (i.e. differences of LD among three planning groups for each task) existed. The ANOVA results indicated that in the simple task, participants' LD did not differ significantly among the OLP (*M* = .509, *SD* = .03), PTP (*M* = .509, *SD* = .03) and comparison groups (*M* = .517, *SD* = .03), *F* (2, 196) = 1.074, *p* = .344. The same was true for participants' LD across the OLP (*M* = .512, *SD* = .03), PTP (*M* = .510, *SD* = .03) and comparison groups (*M* = .506, *SD* = .03) in the complex task, *F*

[13] MD. = -.0004

(2, 196) = .784, *p* = .458. The results indicate different planning types did not affect students' lexical density either in the simple, or in the complex versions of the task.

4.3.2.2 Effects on WT/√2W

The effects of cognitive-task variability and planning type on WT/√2W as a measure of lexical diversity are reported in this section.

Figure 4.9 depicts WT/√2W production of participants under the three planning groups declined when the cognitive-task variability was increased. The decrease was greatest for the CG. The OLP group produced the most WT/√2W in both simple and complex task versions.

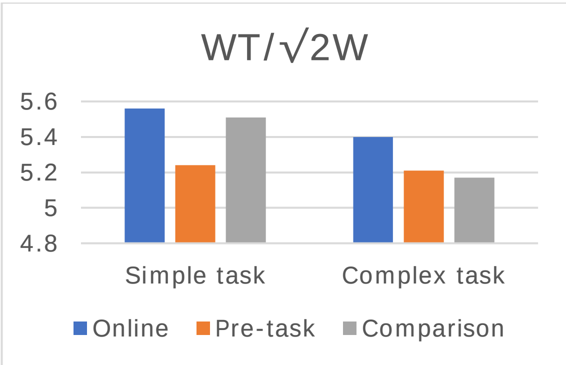

Figure 4.9 Mean WT/√2W for simple and complex tasks across three planning conditions

The assumption of normality was met. Levene's *F* test for the simple task, *F* (2, 198) = .095, *p* = .909, and for the complex task, *F* (2, 198) = .597, *p* = .551, was not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA and further paired-samples *t*-tests and one-way ANOVAs were, therefore, conducted.

Table 4.26 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on WT/√2W

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	31.580	.000***	.138
group	2	198	5.809	.004**	.055
task*group	2	396	8.129	.000***	.076

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

The mixed ANOVA results (Table 4.26) showed that the main effects for cognitive-task variability F (1, 396) = 31.580, p< .001, η_p^2 = .138 and for planning type, F (2, 198) = 5.809, p= .004, η_p^2 = .055 had significant differences. Interactions between cognitive-task variability and planning type in participants' WT/√2W production were also statistically significant, F (2, 396) = 8.129, p < .001 η_p^2 = .076. These results suggest that students' production of WT/√2W is influenced significantly by increasing the cognitive-task variability (without considering the effects of planning type) and by providing different planning types (without considering the effects of cognitive-task variability). Students' WT/√2W production was influenced by the interaction of cognitive-task variability and planning type.

Further follow-up analyses were conducted to examine where the significant within-subjects differences and between-subjects differences located.

Table 4.27 Paired-samples *t*-tests for WT/√2W across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.34	.06	5.756	.000***	.692
OLP (S-C)	68	.17	.06	2.960	.004**	.320
PTP (S-C)	68	.03	.05	.546	.587	.057

Note. S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

The results of paired-samples *t*-tests, as shown in Table 4.27, indicate that the PTP group produced similar WT/√2W in the simple (M = 5.24, SD = .46) and complex tasks, (M =

5.21, *SD* = .48), *t* (67) = .546, *p* = .587, *d* = .057. However, with the CG, the students completing the simple task (*M* = 5.51, *SD* = .51) outperformed those completing the complex task (*M* = 5.17, *SD* = .48) in terms of WT/√2W, *t* (64) = 5.756, *p*<.001, *d* = .692, with a medium to large effect size. Likewise, the OLP group's production of WT/√2W in the complex task, (*M* = 5.40, *SD* = .53), was statistically lower than their performance in the simple task, (*M* = 5.56, *SD* = .52), *t* (67) = 2.960, *p* = .004, *d* = .320, with a small effect size. These results suggest that the CG and OLP conditions inhibited students' production of WT/√2W when the cognitive-task variability was increased. Pre-task planning did not affect students' WT/√2W performance when the cognitive-task variability was increased.

Table 4.28 Post-hoc Comparisons of WT/√2W among Three Planning Type for Two Tasks

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Simple task	CG-OLP	-.05	.09	1.000	-.098
	CG-PTP	.28	.09	.004**	.576
	OLP-PTP	.33	.08	.000***	.672
Complex task	CG-OLP	-.22	.09	.030**	-.444
	CG-PTP	-.04	.09	1.000	-.075
	OLP-PTP	.19	.09	.086	.371

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
***p*<.05; ****p*<.001

One-way ANOVAs, used to analyse the between-group contrasts of WT/√2W in each task version, identified significant differences among the three planning groups when students were under the simple task condition, *F* (2, 198) = 8.668, *p*<.001, and the complex task condition, *F* (2, 198) = 3.930, *p* = .021. Further pairwise comparisons showed (see Table 4.28) that the CG students (*M* = 5.51, *SD* = .51), in the simple task, significantly outperformed the PTP students (*M* = 5.24, *SD* = .46) in WT/√2W production, *p* = .004, *d* = .576, with an effect size of medium to large. Likewise, online planners (*M* = 5.56, *SD* = .52) produced statistically more WT/√2W in the simple task than pre-task planners (*M* = 5.24, *SD* = .46), *p*<.001, *d* = .672 with a medium effect. There was no significant difference in WT/√2W between the CG (*M* = 5.51, *SD* = .51) and the OLP group (*M* = 5.56, *SD* = .52), *p* = 1.000, *d* = -.098. With the complex task, there were no differences between the CG (*M* = 5.17, *SD* = .48) and the PTP group (*M* = 5.21, *SD* = .48) in WT/√2W production, *p* = 1.000, *d* = -.075. Likewise, no significant differences in WT/√2W were found between the OLP (*M* = 5.40, *SD* = .53) and PTP groups (*M* = 5.21, *SD* = .48), *p* = .086, *d* = .371. WT/√2W produced by the OLP group (*M* = 5.40, *SD* = .53), however, markedly exceeded the CG (*M* = 5.17, *SD* = .48), *p* = .030, *d* = -.444, with an effect size close to medium. The results reveal that when cognitive task complexity was at a low level, the OLP group and the CG outperformed the PTP group in terms of WT/√2W. When cognitive task complexity was at a high level, the OLP group produced significantly more WT/√2W than the CG.

The interaction graph (see Figure 4.9) reveals that participants' WT/√2W production under the OLP and CG conditions decreased significantly as cognitive-task variability was increased, which was more pronounced for the CG, suggesting the CG was more influenced by the increase of cognitive-task variability than the OLP group. For the PTP group, there was no significant difference in WT/√2W between the simple and complex tasks, although the WT/√2W performance slightly declined when the cognitive-task variability was more complex.

4.3.2.3 Summary

In this section, the effects of cognitive-task variability and planning type on lexical complexity, including lexical density and lexical diversity, are summarized (see Table 4.29).

The within-subjects differences, as an outcome of cognitive-task variability, are summarised as the follows.

- No significant differences in lexical density occurred in the OLP and PTP groups, while the CG students' LD under the simple task was greater than with the complex task.
- Increasing cognitive task complexity led to a significant reduction in participants' lexical diversity (WT/√2W) under the OLP and CG conditions, but there were no obvious changes under the PTP condition.

Table 4.29 Summary of Effects of Cognitive-task Variability and Planning Type on Lexical Complexity

Lexical complexity measures	Within-subjects differences (S vs. C)			Between-subjects differences (OLP vs. PTP vs. CG)		Interaction effects
	OLP	PTP	CG	simple task	complex task	
LD	no Sig.	no Sig.	S>C	no Sig.	no Sig.	no Sig.
WT/√2W	S>C	no Sig.	S>C	OLP>PTP CG>PTP	OLP>CG	Sig.

Note. S = Simple task, C = Complex task, OLP = online planning group, PTP = pre-task planning group, CG = comparison group, no Sig. = no significant difference.

The between-group differences, related to the impact of planning type, can be summarised in the following aspects.

- For the lexical density, neither interaction effects nor between-subjects differences were found.
- With the simple task, lexical diversity (WT/√2W) of both the CG and OLP groups was greater than that of the PTP group. With the complex task, the OLP group produced more WT/√2W than the CG. The interaction effect of cognitive-task variability and planning type on WT/√2W also identified that the CG was more negatively influenced than the OLP group by manipulating the cognitive-task variability.

4.3.3 Effects on accuracy

This section reports the findings of the effects of cognitive-task variability and planning type on students' accuracy. Err/T and EP100 were used to identify the presence, and EFC/C to the absence of errors. Table 4.30 displays the descriptive results for the seven syntactic complexity measures in the simple and complex tasks across the CG, OLP and PTP groups.

Table 4.30 Descriptive Statistics for Accuracy across Three Planning Groups in Two Tasks

Syntactic complexity	Groups	*N*	Simple task		Complex task	
			M	*SD*	*M*	*SD*
Err/T	CG	65	.66	.33	.58	.27
	OLP	68	.74	.29	.73	.37
	PTP	68	.80	.28	.73	.34
EP100	CG	65	.053[14]	.02	.050	.02
	OLP	68	.060	.02	.061	.03
	PTP	68	.068	.03	.062	.03
EFC/C	CG	65	.66	.14	.67	.12
	OLP	68	.63	.12	.63	.15
	PTP	68	.59	.15	.61	.20

Note. PTP=pre-task planning group, OLP=online planning group, CG=comparison group, Err/T=the number of errors per T-unit, EP100=the number of errors per 100 words, EFC/C= ratio of error-free clauses.

4.3.3.1 Effects on Err/T

Figure 4.10 depicts the mean scores of Err/T for participants under the CG, OLP and PTP conditions in the simple and complex tasks. According to the graph, the PTP and CG saw a decreasing trend, and there were hardly changes in the OLP group when cognitive-task variability was increased. In both simple and complex tasks, the PTP group made the most and the CG the fewest Err/T.

[14] Mean here was rounded to 3 decimal places, since the numbers are too close.

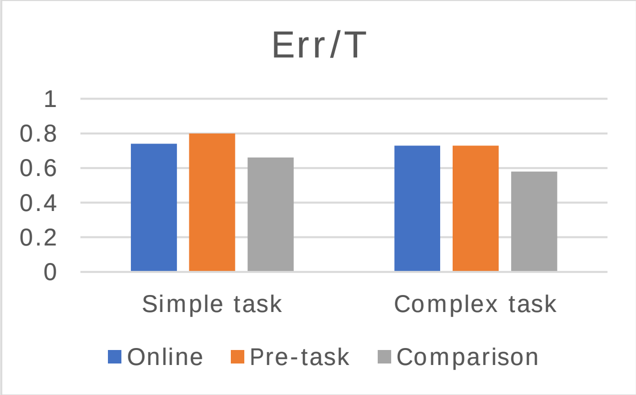

Figure 4.10 Mean Err/T for simple and complex tasks across three planning conditions

The assumption of normal distribution was first checked and met. Levene's F test for the simple task, F (2, 198) = .661, p = .518, and for the complex task, F (2, 198) = 1.602, p = .204, showed that the assumption of homogeneity of the variances was not violated for either the simple or complex task. Hence, a mixed ANOVA was conducted to explore the effects of cognitive-task variability and planning type on students' Err/T.

Table 4.31 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on Err/T

	df1	*df2*	*F*	Sig.	η_p^2
task	1	396	5.614	.019**	.028
group	2	198	5.132	.007**	.049
task*group	2	396	.576	.563	.006

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Table 4.31 showed, the mixed ANOVA yielded a significant main effect on participants' Err/T for cognitive-task variability, F (1, 396) = 5.614, p = .019, η_p^2 = .028, as well as for planning type, F (2, 198) = 5.132, p = .007, η_p^2 = .049. No significant differences were found, however, for the interaction between cognitive-task variability and planning type F (2, 396) = .576, p = .563, η_p^2 = .006 in participants' Err/T production. These results indicate that increasing cognitive-task variability (if the planning type was ignored), and providing different planning types (if the cognitive-task variability was ignored) significantly affected students' accuracy. Students' accuracy was, however, not interactively influenced by cognitive-task variability and planning type.

Three separate paired-samples *t*-tests were conducted to test where the within-subjects differences were located concerning participants' Err/T under the OLP, PTP and CG conditions respectively.

Table 4.32 Paired-samples *t*-tests for Err/T across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.08	.04	2.123	.038**	.247
OLP (S-C)	68	.02	.04	.449	.665	.060
PTP (S-C)	68	.07	.04	1.761	.083	.217

Note. Err/T = The number of errors per T-unit, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
$^{**}p<.05$; $^{***}p<.001$

Table 4.32 shows that in the comparison group, participants' Err/T differed significantly in the simple task (M = .66, SD = .33) compared to Err/T in the complex task (M = .58, SD = .27), t (64) = 2.123, p = .038, d = .247. The effect size was small. No significant differences in Err/T were identified for the PTP group between the simple (M = .80, SD = .28) and complex tasks (M = .73, SD = .34), t (67) = 1.761, p = .083, d = .217, with a small effect size. Likewise, the OLP group made similar Err/T between the simple and complex tasks, t (67) = .449, p = .665, d = .060. These findings suggest that increasing cognitive-task variability helped participants make fewer Err/T under the PTP and CG conditions, but it had no significant effects for the OLP group.

Two one-way ANOVA tests were conducted to examine the differences of Err/T across three planning groups under both the simple task and the complex task conditions. The ANOVA results showed that the between-subjects differences among the three planning groups were significant for both the simple task, F = 3.568, p = .030, and the complex task, F = 4.222, p = .016.

Table 4.33 Post-hoc Comparisons of Err/T among Three Planning Type for Two Tasks

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Simple task	CG-OLP	-.09	.05	.313	-.273
	CG-PTP	-.14	.05	.026**	-.456
	OLP-PTP	-.05	.05	.911	-.186
Complex task	CG-OLP	-.14	.06	.043**	-.435
	CG-PTP	-.15	.06	.032**	-.479
	OLP-PTP	-.01	.06	1.000	-.010

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
$^{**}p<.05$; $^{***}p<.001$

Pairwise comparisons showed that (see Table 4.33) in the simple task condition, Err/T made by the PTP group (M = .80, SD = .28) outnumbered those made by the CG (M = .66, SD = .33), p = .026, d = -.456 with a medium effect size. No significant differences in Err/T were found, however, between the CG (M = .66, SD = .33) and OLP groups (M = .74, SD = .29), p = .313, d = -.273, and between the OLP (M = .74, SD = .29) and PTP groups (M = .80, SD = .28), p = .911, d = -.186. Whereas in the complex version of the task, the CG (M = .58, SD = .27) made significantly fewer Err/T than the OLP (M = .73, SD = .37), p = .043, d = -.435, and PTP groups (M = .73, SD = .34), p = .032, d = -.479, and both effect sizes were medium; there were no significant contrasts in Err/T between the OLP (M = .73, SD = .37) and PTP groups (M = .73, SD = .34), p = 1.000, d = -.010. These results reveal that, in both the simple and complex tasks, the CG achieved the highest accuracy with the fewest Err/T among the three planning types.

4.3.3.2 Effects on EP100

In this section, the results for the effects of cognitive-task variability and planning type on EP100 as a measure of accuracy are reported. As Figure 4.11 illustrates, EP100 under the PTP and CG conditions decreased when cognitive-task variability increased. There were few changes of EP100 made by the OLP group under different cognitive-task variability levels.

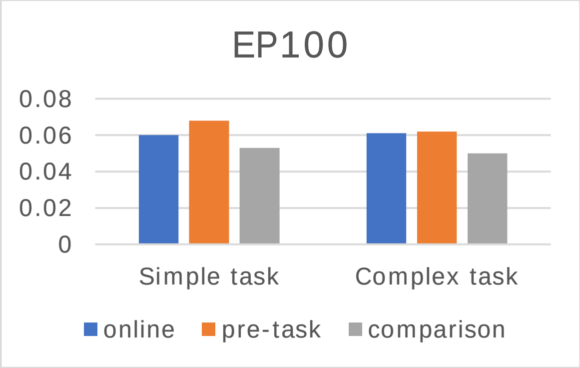

Figure 4.11 Mean EP100 for simple and complex tasks across three planning conditions

The assumption of normality was met, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* tests in the simple task, *F* (2, 198) = .741, *p* = .478, and in the complex task, *F* (2, 198) = 1.136, *p* = .323). A mixed ANOVA, therefore, was computed.

Table 4.34 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on EP100

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	3.006	.085	.015
group	2	198	6.878	.001**	.065
task*group	2	396	1.530	.219	.015

Note. Task = cognitive-task variability, group = planning type.
***p*<.05; ****p*<.001

The mixed ANOVA (see Table 4.34) yielded no significant differences in EP100 for either the main effect of cognitive-task variability, *F* (1, 396) = 3.006, *p* = .085, η_p^2 = .015, or the interaction between cognitive-task variability and planning type, *F* (2, 396) = 1.530, *p* = .219, η_p^2 = .015. The mixed ANOVA, however, detected a significant main effect for planning type *F* (2, 198) = 6.878, *p* = .001, η_p^2 = .065. These results indicate that varying the planning types significantly influenced participants' accuracy in terms of EP100 (without considering the effects of cognitive-task variability). Students' accuracy was, however, not affected by cognitive-task variability (without considering the effects of planning type), or by the interaction between cognitive-task variability and planning type.

Three separate paired-samples *t*-tests were performed to test the within-subjects differences in participants' accuracy in terms of EP100 under the CG, OLP and PTP conditions respectively.

Table 4.35 Paired-samples *t*-tests for EP100 across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.004	.003	1.235	.221	.159
OLP (S-C)	68	-.001	.003	-.318	.752	-.038
PTP (S-C)	68	.006	.003	2.150	.035**	.238

Note. S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**$p<.05$; ***$p<.001$

Table 4.35 shows that neither under the CG condition, t (64) = 1.235, p = .221, d = .159, nor the OLP condition, t (67) = -.318, p = .752, d = -.038, did participants' EP100 differ statistically between the simple task (CG: M = .053, SD = .02; OLP: M = .060, SD = .02) and the complex task (CG: M = .050, SD = .02; OLP: M = .061, SD = .03). The PTP students made significantly more EP100 in the simple task (M = .068, SD = .03) than in the complex one (M = .062, SD = .03), t (67) = 2.150, p = .035, d = .238, and the effect size was small. These findings reveal that increasing cognitive-task variability did not significantly affect EP100 for students who were in the CG and OLP groups, but helped students write more accurate essays in EFL writing under the PTP condition.

Table 4.36 Post-hoc Comparisons of EP100 among Three Planning Type for Two Tasks

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Simple task	CG-OLP	-.007	.004	.269	-.307
	CG-PTP	-.015	.004	.002**	-.598
	OLP-PTP	-.008	.004	.205	-.309
Complex task	CG-OLP	-.012	.004	.022**	-.477
	CG-PTP	-.012	.004	.016**	-.512
	OLP-PTP	-.001	.004	1.000	-.043

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
***p*<.05; ****p*<.001

Two one-way ANOVA tests were conducted to examine the differences of EP100 across the three planning groups with both the simple and complex tasks. The ANOVA results showed that for the simple task, the between-subjects differences among the three planning groups were significant, F (2,198) = 6.191, p = .002. Likewise, a significant difference for between-group contrasts was found under the complex task condition, F (2,198) = 5.070, p = .007.

Pairwise comparisons with Bonferroni adjustment were further applied; the results are reported in Table 4.36. As the table displays, in the simple task, no significant differences were found between the CG (M = .053, SD = .02) and OLP groups (M = .060, SD = .02), p = .269, d = -.307, or between the OLP (M = .060, SD = .02) and PTP groups (M = .068, SD = .03), p = .205, d = -.309; significant differences were identified between the CG (M = .053, SD = .02) and PTP groups (M = .068, SD = .03), p = .002, d = -.598, with a close to large effect size. Whereas in the complex task, online planners (M = .061, SD = .03) made similar EP100 to pre-task planners (M = .062, SD = .03), p = 1.000, d = -.043, the EP100 made by the OLP group (M = .061, SD = .03) substantially outnumbered those by the CG (M = .050, SD = .02), p = .022, d = -.477, with a medium effect size. The PTP group (M = .062, SD = .03) also made significantly more EP100 than the CG (M = .050, SD = .02), p = .016, d = -.512, with medium to large effect size. These results suggest that the CG condition, compared with the OLP and PTP group, had a beneficial impact on students' accuracy in both the simple and complex tasks.

4.3.3.3 Effects on EFC/C

Figure 4.12 illustrates the changes of EFC/C across the CG, OLP and PTP groups from the simple task to the complex task. The graph shows that EFC/C of the CG and PTP groups was higher in the complex task than the simple task. Hardly changes in EFC/C were detected for the OLP group. In both simple and complex tasks, the most accurate essays were written by the CG, followed by the OLP group; the PTP group made the least EFC/C among the three groups but was still higher in the complex task than the simple task.

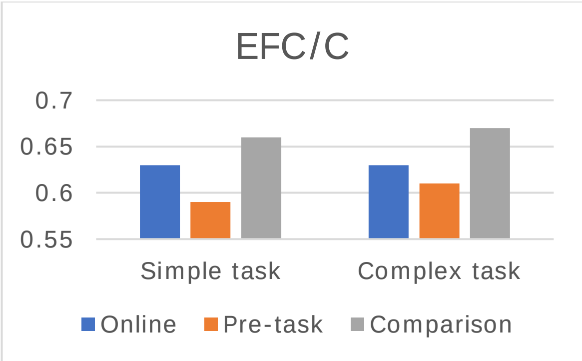

Figure 4.12 Mean EFC/C for simple and complex tasks across three planning conditions

To test the impact of cognitive-task variability and planning type on EFC/C, first, the normal distribution of participants' score was evaluated and found to be met. A decision was initially made to perform a mixed ANOVA. The homogeneity of variances assumption was, however, violated for the complex task, as Levene's F test in the complex task, F (2, 198) = 3.637, p = .028, was significant, although Levene's F test was not significant in the simple task, F (2, 198) = .952, p = .388. As such, other appropriate tests (i.e., paired-sample t-tests and one-way ANOVAs) instead of mixed ANOVA were conducted.

Three paired-samples t-tests were conducted to explore where the within-subjects differences in learners' EFC/C were located under the OLP, PTP and CG conditions respectively.

Table 4.37 Paired-samples *t*-tests for EFC/C across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	-.02	.02	-.861	.392	-.112
OLP (S-C)	68	.00[15]	.02	.105	.917	.014
PTP (S-C)	68	-.03	.02	-1.242	.218	-.159

Note. EFC/C = Ratio of error-free clauses, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std. = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
***p*<.05; ****p*<.001

Table 4.37 shows that in the CG, participants' EFC/C was similar in the simple task (M = .66, SD = .14), compared to that in the complex task (M = .67, SD = .12), t (64) = -.861, p = .392, d = -.112. The same pattern was also found for the OLP, t (67) = .105, p = .917, d = .014, and PTP groups, t (67) = -1.242, p = .218, d = -.159. These findings suggest that cognitive-task variability did not significantly affect students' EFC/C performance under the CG, OLP, and PTP conditions.

Two one-way ANOVA tests were conducted to explore the differences in EFC/C across the three planning groups under the simple task and the complex task conditions. The ANOVA results showed that the between-subjects differences among the three planning groups were statistically significant for the simple task condition, F (2,198) = 4.191, p = .016. The Robust tests[16] showed that no significant differences in EFC/C across the three planning groups when students completed the complex task, F (2,173.64) = 2.411, p = .093.

[15] = .002
[16] Brown-Forsythe test in the table Robust Tests of Equality of Means.

Table 4.38 Post-hoc Comparisons of EFC/C among Three Planning Type for the Simple Task

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Simple	CG-OLP	.03	.02	.717	.217
task	CG-PTP	.07	.02	.013**	.467
	OLP-PTP	.04	.02	.264	.299

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

Pairwise comparisons presented in Table 4.38 show that with the simple version of the task, the CG (M = .66, SD = .14) made significantly more EFC/C than the PTP group (M = .59, SD = .15), p = .013, d = .467, with a medium effect size. However, there was no significant difference in EFC/C between the CG (M = .66, SD = .14) and OLP groups (M = .63, SD = .12), p = .717, d = .217, or between the OLP and PTP groups (M = .59, SD = .15), p = .264, d = .299. These results indicate that the CG's accuracy was better than the PTP group with the simple task, although no significant differences in EFC/C were found across the three groups under the complex task condition.

4.3.3.4 Summary

In this section, the effects of cognitive-task variability and planning type on students' accuracy in terms of Err/T, EP100 and EFC/C are summarized (see Table 4.39).

The within-subjects differences for the effects of cognitive-task variability are summarised as following. Increasing cognitive-task variability led to significantly greater students' accuracy under both PTP (EP100 as measures) and CG conditions (Err/T as a measure). No obvious changes in students' accuracy were identified, as cognitive-task variability was increased under the OLP condition.

Planning type had some effects in the following aspects, as evident in the between-subjects differences. The PTP group made more errors than the CG under the simple task condition with students' accuracy measured by Err/T and EP100. Fewer errors were identified in the CG condition, compared to the PTP and OLP conditions when the task was cognitively demanding. As measured by EFC/C, the CG had higher accuracy scores than the PTP group under the simple task condition, while there were no significant differences across the three planning groups when participants completed the complex task.

Table 4.39 Summary of Effects of Cognitive-task Variability and Planning Type on Accuracy

Accuracy measures	*Within-subjects differences (S vs. C)*			*Between-subjects differences (OLP vs. PTP vs. CG)*		*Interaction effects*
	OLP	*PTP*	*CG*	*simple task*	*complex task*	
Err/T	no Sig.	no Sig.	S>C	PTP >CG	PTP>CG OLP>CG	no Sig.
EP100	no Sig.	S>C	no Sig.	PTP >CG	PTP>CG OLP>CG	no Sig.
EFC/C	no Sig.	no Sig.	no Sig.	CG>PTP	no Sig.	-

Note. S = Simple task, C = Complex task, OLP = online planning group, PTP = pre-task planning group, CG = comparison group, no Sig. = no significant difference.

4.3.4 Effects on fluency

Participants' writing fluency was measured in the present study to capture the effects of cognitive-task variability and planning type on students' fluency. Table 4.40 below presents the descriptive statistics for students' writing fluency in the simple and the complex tasks across the CG, OLP and PTP groups.

Table 4.40 Descriptive Statistics for Fluency across Three Planning Groups in Two Tasks

Groups	*N*	Simple task		Complex task	
		M	*SD*	*M*	*SD*
CG	65	5.54	.93	5.48	.86
OLP	68	5.88	.92	5.89	1.04
PTP	68	5.55	.93	6.05	.94

Note. PTP=pre-task planning group, OLP=online planning group, CG=comparison group.

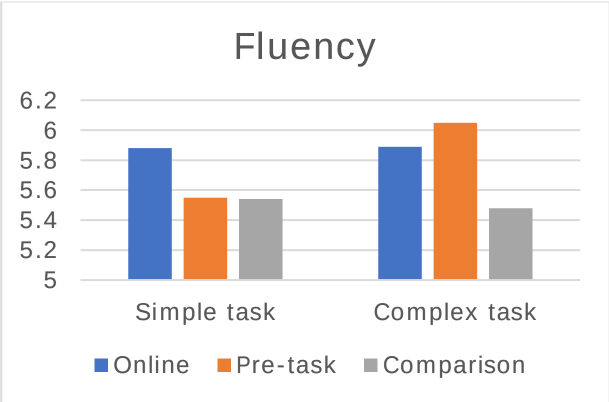

Figure 4.13 Mean fluency for simple and complex tasks across three planning conditions

In Figure 4.13, the mean scores for fluency presented in Table 4.40 illustrates that the PTP participants' fluency was higher in the complex task than the simple task, while there was little difference in OLP and CG participants' fluency due to the increases in cognitive-task variability. Furthermore, in the complex task, the fluency of the PTP group was greater than that of the OLP and comparison groups. Among the three groups, the CG students' fluency was the lowest in both simple and complex task.

The assumption of normal distribution was first evaluated and found to be satisfactory. The assumption of homogeneity of variances was not violated, with Levenes' *F* test not being significant for either the simple, F (2, 198) = .202, p = .818 or the complex task, F (2, 198) = 1.431, p = .242. Hence, a decision was made to perform a mixed ANOVA.

Table 4.41 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on Fluency

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	5.402	.021**	.027
group	2	198	3.679	.027**	.036
task*group	2	396	7.792	.001**	.073

Note. Task = cognitive-task variability, group = planning type.
$**p<.05$; $***p<.001$

As Table 4.42 shows, the mixed ANOVA yielded a significant main effect on participants' fluency for cognitive-task variability, F (1, 396) = 5.402, p = .021, η_p^2 = .027, and for planning type, F (2, 198) = 3.679, p = .027, η_p^2 = .036. Significant differences were found for the interaction between cognitive-task variability and planning type F (2, 396) = 7.792, p = .001, η_p^2 = .073 in terms of participants' fluency. These results suggest that cognitive-task variability and planning type interacted to influence students' fluency.

Table 4.42 Paired-samples *t*-tests for Fluency across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.06	.13	.466	.643	.065
OLP (S-C)	68	-.00[17]	.12	-.018	.985	-.002
PTP (S-C)	68	-.50	.08	-6.407	.000***	-.529

Note. FL = writing fluency, S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
$**p<.05$; $***p<.001$

The three paired-samples *t*-tests were conducted to examine the within-subjects differences across the CG, OLP and PTP groups. As shown in Table 4.42, students' fluency under the CG and OLP conditions was not significantly different between their simple task (CG: M = 5.54, SD = .93; OLP: M = 5.88, SD = .92) and their complex task completion (CG: M = 5.48, SD = .86; OLP: M = 5.89, SD = 1.04), t (64) = .466, p = .643, d = .065; t (67) = -.018, p = .985, d = -.002. The PTP students performed better in the

[17] = -.002

complex task (*M* = 6.05, *SD* = .94) than in the simple task (*M* = 5.55, *SD* = .93), *t* (65) = -6.407, *p*<.001, *d* = -.529, with a medium effect size.

Two separate one-way ANOVA tests were conducted to examine whether significant between-subjects differences existed. The ANOVA results showed that the between-subjects differences among the three planning groups were not significant for the simple task condition, *F* = 2.993, *p* = .052, but were significant for the complex task, *F* = 6.146, *p* = .003.

Table 4.43 Post-hoc Comparisons of Fluency among Three Planning Type for the Complex Task

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Complex task	CG-OLP	-.40	.16	.046**	-.422
	CG-PTP	-.56	.16	.002**	-.624
	OLP-PTP	-.16	.16	.991	-.161

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
***p*<.05; ****p*<.001

Pairwise comparisons showed (see Table 4.43), as for the complex task, students in the CG (*M* = 5.48, *SD* = .86) wrote significantly less fluently than their counterparts in the PTP (*M* = 6.05, *SD* = .94), *p*=.002, *d* = -.624 and OLP groups (*M* = 5.89, *SD* = 1.04), *p*=.046, *d* = -.422. Both effect sizes were medium. The fluency of the OLP group (*M* = 5.89, *SD* = 1.04) was similar to that of the PTP group (*M* = 6.05, *SD* = .94), *p* = .991, *d* = -.161. These results indicate that pre-task planning and online planning had a more positive effect on students' writing fluency in the complex task than for the comparison group. No significant differences were, however, found for fluency among the three groups in the simple condition.

The interaction graph (see Figure 4.13) reveals that students' fluency under the PTP condition significantly increased as cognitive-task variability was increased.

4.3.4.1 Summary

As summarised in Table 4.44, increases in cognitive-task variability had a beneficial impact on PTP students’ fluency. Neither the OLP nor CG conditions appeared influenced by the increased cognitive-task variability.

The between-subjects differences indicate that the PTP and OLP groups wrote significantly more fluently than the CG group in the complex task. No significant between-group differences in fluency were found under the simple task condition.

Table 4.44 Summary of Effects of Cognitive-task Variability and Planning Type on Fluency

Fluency measures	*Within-subjects differences (S vs. C)*			*Between-subjects differences (OLP vs. PTP vs. CG)*		*Interaction effects*
	OLP	*PTP*	*CG*	*simple task*	*complex task*	
Fluency	no Sig.	S<C	no Sig.	no Sig.	PTP>CG OLP>CG	Sig.

Note. S = Simple task, C = Complex task, OLP = online planning group, PTP = pre-task planning group, CG = comparison group, no Sig. = no significant difference.

4.3.5 Effects on adequacy

This section reports the findings for the effects of cognitive-task variability and planning type on students’ adequacy. Three measures were used: subjective rating on content, organization and overall text quality. Jacobs et al.’s (1981) analytical rating scheme was employed to quantify these components. Table 4.45 displays the descriptive results for adequacy measures in the simple and complex tasks across the CG, OLP and PTP groups.

Table 4.45 Descriptive Statistics for Adequacy across Three Planning Groups in Two Tasks

Syntactic complexity	Groups	*N*	Simple task		Complex task	
			M	*SD*	*M*	*SD*
Content	CG	65	18.56	1.77	18.16	1.53
	OLP	68	18.55	1.44	18.03	1.38
	PTP	68	18.45	1.48	18.92	1.47
Organisation	CG	65	14.23	1.63	13.85	1.66
	OLP	68	13.99	1.28	13.60	1.30
	PTP	68	14.10	1.57	14.46	1.55
Overall writing score	CG	65	64.83	5.13	63.10	5.50
	OLP	68	64.38	4.05	63.14	4.48
	PTP	68	63.18	5.25	64.24	5.25

Note. PTP=pre-task planning group, OLP=online planning group, CG=comparison group.

4.3.5.1 *Effects on content*

Figure 4.14 reflects the changes of students' mean scores for content. As the graph illustrates, with the increase of cognitive-task variability, students' content scores under the OLP and CG conditions decreased, while under the PTP condition students' content scores increased.

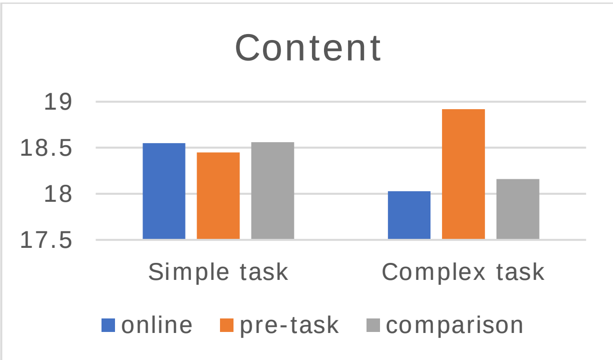

Figure 4.14 Mean content for simple and complex tasks across three planning conditions

The assumption of normality was tenable. Levene's F tests for content score in the simple task, F (2, 198) = .884, p = .415, and for content score in the complex task, F (2, 198) = .644, p = .526, were not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed

ANOVA was, therefore, conducted to assess the impact of cognitive-task variability and planning type on participants' score of content.

Table 4.46 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on Content

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	2.180	.141	.011
group	2	198	1.681	.189	.017
task*group	2	396	9.501	.000**	.088

Note. Task = cognitive-task variability, group = planning type, η_p^2 = partial η^2.
**p<.05; ***p<.001

The results of the mixed ANOVA in Table 4.46 showed that a substantial significant interaction between task and group was identified, F (2, 396) = 9.501, p < .001, η_p^2 = .088, revealing that students' writing content in the simple and the complex versions of the tasks was contingent upon the planning conditions. The main effect for cognitive-task variability was, however, not significant, F (1, 396) = 2.180, p = .141, η_p^2 = .011, suggesting no difference as a result of cognitive-task variability. Similarly, no significant main effect for planning type was found, F (2, 198) = 1.681, p = .189, η_p^2 = .017, indicating no significant variance in students' content score across the three groups.

To further examine the interaction effects, followed-up paired t-tests were applied to compare within-subjects differences across the CG, OLP, and PTP groups.

Table 4.47 Paired-samples *t*-tests for Content across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.40	.19	2.099	.040*	.240
OLP (S-C)	68	.52	.16	3.341	.001*	.370
PTP (S-C)	68	-.47	.18	-2.585	.012*	-.319

Note. S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

The results as shown in Table 4.47 revealed that in the CG, students' content in the simple task (M = 18.56, SD = 1.77) exceeded that in the complex task (M = 18.16, SD = 1.53), t (64) = 2.099, p = .040, d = .240, although the effect size was small. Similarly, the content of the online planners was better in the simple task (M = 18.55, SD = 1.44) than the complex one (M = 18.03, SD = 1.38), t (67) = 3.341, p = .001, d = .370, also with a small effect size. By contrast, for the pre-task planners, content score significantly increased, t (67) = -2.585, p = .012, d = -.319, from the simple (M = 18.45, SD = 1.48) to the complex task versions (M = 18.92, SD = 1.47). The findings suggest that increasing cognitive-task variability had statistically negative effects on students' content under the CG and OLP conditions, but a statistically positive impact on the PTP group.

Table 4.48 Post-hoc Comparisons of Content among Three Planning Type for the Complex Tasks

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Complex task	CG-OLP	.13	.25	1.000	.09
	CG-PTP	-.76	.25	.009**	-.505
	OLP-PTP	-.89	.25	.001**	-.623

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

Two separate one-way ANOVA tests were conducted to examine whether there were significant between-subjects differences in content. The ANOVA results showed that students' content did not differ significantly among the OLP (M = 18.55, SD = 1.44), PTP (M = 18.45, SD = 1.48) and comparison groups (M = 18.56, SD = 1.77), in the simple task, F (2, 198) = .107, p = .898. In contrast, significant differences were identified among the three groups in the complex task, F (2, 198) = 7.309, p = .001. As shown in Table 4.48, the PTP group (M = 18.92, SD = 1.47) significantly outperformed both the CG (M = 18.16, SD = 1.53), p = .009, d = -.505 and OLP groups (M = 18.03, SD = 1.38), p = .001, d = -.623 in content with the complex task; both effect sizes were medium to large. These results indicate that with the complex task, pre-task planning had a more positive effect on students' writing content than online planning and comparison group.

As for the interaction effects, the graph (see Figure 4.14) reveals that students' content in their essays benefited the most from the pre-task planning condition and increased cognitive-task variability.

4.3.5.2 Effects on organization

Figure 4.15 illustrates the descriptive statistics of organisation presented in Table 4.45. According to the graph, the organisation scores for the PTP group were higher in the complex task than the simple, whereas the scores of the CG and OLP groups reduced as a function of increasing cognitive-task variability. In both the simple and complex tasks, the OLP group got the lower score for organisation than the PTP and CG groups.

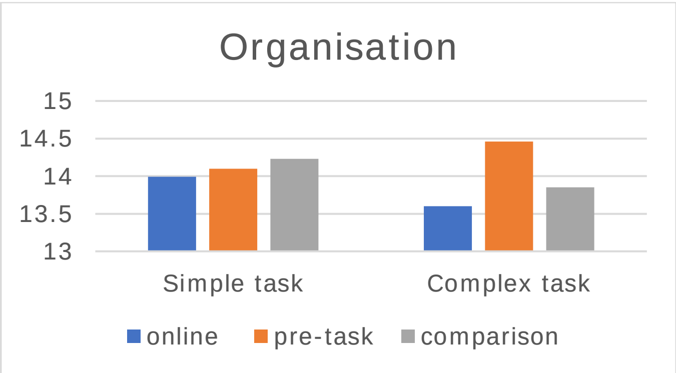

Figure 4.15 Mean organisation for simple and complex tasks across three planning conditions

The assumption of normal distribution of organisation was first checked and met. Levene's F test for the simple task, F (2, 198) = 2.804, p = .063, and for the complex task, F (2, 198) = 2.558, p = .080, showed that the assumption of homogeneity of variances was not violated. Hence, a mixed ANOVA was conducted to explore the effects of cognitive-task variability and planning type on students' organisation performance.

Table 4.49 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on Organisation

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	1.637	.202	.008
group	2	198	2.292	.104	.023
task*group	2	396	5.849	.003**	.056

Note. Task = cognitive-task variability, group = planning type, η_p^2 = partial η^2.
**p<.05; ***p<.001

The mixed ANOVA was applied to calculate the effects of cognitive-task variability and planning type on organisation; the results are shown in Table 4.49. The interaction between cognitive-task variability and planning type, F (2, 396) = 5.849, p = .003, η_p^2 = .056 was statistically significant with a medium effect size, which indicates that students' writing organisation on the simple versus the complex versions of the tasks was contingent upon the planning conditions. Nevertheless, significant main effects were not found for cognitive-task variability, F (1, 396) = 1.637, p = .202, η_p^2 = .008, or for planning type F (2, 198) = 2.292, p = .104, η_p^2 = .023. The results reveal that neither varying cognitive-task variability (if the planning type was ignored) nor providing different planning conditions (if the cognitive-task variability was ignored) affected students' organisation significantly.

To further investigate the interaction effects, followed-up paired samples t-tests were computed to compare within-subjects differences across the CG, OLP and PTP groups. The findings of three paired-samples t-tests are reported in Table 4.50.

Table 4.50 Paired-samples *t*-tests for Organisation across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	.38	.19	2.002	.050**	.229
OLP (S-C)	68	.39	.18	2.224	.030**	.302
PTP (S-C)	68	-.37	.18	-2.080	.041**	-.236

Note. S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

CG students' organisation score in the simple task (M = 14.23, SD = 1.63) was higher than that in the complex task (M = 13.85, SD = 1.66), t (64) = 2.002, p = .050, d = .229; the effect size was small. The same pattern was also found in the OLP group with the simple task (M = 13.99, SD = 1.28), compared with the complex task versions (M = 13.60, SD = 1.30), t (67) = 2.224, p = .030, d = .302, and the effect size was small. A substantial increase of the organisation score occurred in the PTP group when the cognitive-task variability was raised from simple (M = 14.10, SD = 1.57) to complex (M = 14.46, SD = 1.55), t (67) = -2.080, p = .041, d = -.236, although the effect size was small. These findings suggest that increasing cognitive-task variability benefited students' organisation performance when they used pre-task planning, but reduced students' organisation score when they were under either the online planning condition, or the comparison group.

Two one-way ANOVA tests were conducted to examine the differences in organisation performance across the three planning groups under the simple task and the complex task conditions. The ANOVA showed that the between-subjects differences among the three planning groups were not significant for the simple task condition, F (2,198) = .421, p = .657, but for the complex task version, F (2,198) = 5.836, p = .003.

Table 4.51 Post-hoc Comparisons of Organisation among Three Planning Type for the Complex Task

task	Group(I)----(J)	MD.(I-J)	Std.Er	Sig.	*d*
Complex task	CG-OLP	.25	.26	1.000	.168
	CG-PTP	-.61	.26	.063	-.376
	OLP-PTP	-.86	.26	.003**	-.602

Note. MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**p<.05; ***p<.001

Pairwise comparisons showed (see Table 4.51) that, when completing the complex task, students under the PTP condition (M = 14.46, SD = 1.55) achieved a significantly higher score of organisation than their counterparts under the OLP condition (M = 13.60, SD = 1.30), p = .003, d = -.602 with a medium effect size. No significant difference in score of organisation was found between the CG (M = 13.85, SD = 1.66) and OLP groups (M = 13.60, SD = 1.30), p = 1.000, d = .168, and between the CG (M = 13.85, SD = 1.66) and PTP groups (M = 14.46, SD = 1.55), p = .063, d = -.376. These results indicate that students under the PTP condition achieved higher scores for organisation in writing than those under the OLP condition, although when they completed the simple task, no substantial differences in students' organisation performance were found.

According to the visual representation of the means presented in Figure 4.15, participants' organisation was enhanced as a function of increasing cognitive-task variability and providing pre-task planning.

4.3.5.3 Effects on overall text quality

In this section, the results for the effects of cognitive-task variability and planning type on overall text quality are reported. Figure 4.16 illustrates the mean scores of overall text quality in the simple and complex tasks over the CG, OLP and PTP groups. As the graph shows, the overall score under the OLP and CG conditions decreased when cognitive-task variability was increased, in contrast to PTP students' overall score increased from the simple task to the complex task.

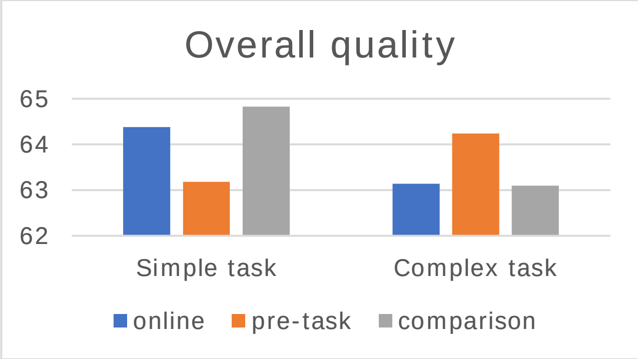

Figure 4.16 Mean overall text quality for simple and complex tasks across three planning conditions

The assumption of normality was met and the assumption of homogeneity of variances was not violated for either the simple or complex task (Levene's *F* tests in both the simple task, *F* (2, 198) = 2.508, *p* = .084, and the complex task, *F* (2, 198) = 1.799, *p* = .168). A mixed ANOVA was, therefore, conducted.

Table 4.52 Mixed ANOVA for Effects of Cognitive-task Variability and Planning Type on Overall Text Quality

	df1	*df2*	*F*	*Sig.*	η_p^2
task	1	396	4.352	.038**	.022
group	2	198	.062	.939	.001
task*group	2	396	7.974	.000**	.075

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

As displayed in Table 4.52, the mixed ANOVA detected a significant interaction effect between cognitive-task variability and planning type, F (2, 396) = 7.974, p < .001, η_p^2 = .075 with a medium effect size and a significant main effect for cognitive-task variability, F (1, 396) = 4.352, p = .038, η_p^2 = .022 with a small effect size. No main effect on students' overall writing quality was, however, found for planning groups, F (2, 198) = .062, p = .939, η_p^2 = .001. These results indicate that students' writing quality on the simple versus the complex task versions was contingent upon planning conditions; providing different planning types alone did not vary students' overall text quality.

To further explore the nature of students' writing quality across the six conditions, the mixed ANOVA analyses were followed by paired-samples *t*-tests for within-groups and one-way ANOVAs for between-groups conditions.

Table 4.53 Paired-samples *t*-tests for Overall Text Quality across Three Planning Groups

	N	*MD.*	*Std.Er*	*t*	*Sig.(2-tailed)*	*d*
CG (S-C)	65	1.73	.54	3.204	.002**	.325
OLP (S-C)	68	1.24	.50	2.459	.017**	.290
PTP (S-C)	68	-1.06	.54	-1.967	.053	-.223

Note. S-C = the score on the simple task-the score on the complex task, MD. = Mean differences, Std = Standard, Er. = Error, CG = comparison group, OLP = online planning group, PTP = pre-task planning group.
**$p<.05$; ***$p<.001$

Table 4.53 reveals that participants' writing quality was significantly lower in the simple task (OLP: M = 64.38, SD = 4.05, CG: M = 64.83, SD = 5.13), than the complex task (OLP: M = 63.14, SD = 4.48, CG: M = 63.10, SD = 5.50), when students were under the OLP, t (67) = 2.459, p = .017, d = .290, and CG conditions, t (64) = 3.204, p = .002, d = .325; both effect sizes were small. By contrast, the PTP students' overall writing score in the complex task (M = 64.24, SD = 5.25) was marginally higher than their score in the simple task (M = 63.18, SD = 5.25), t (67) = -1.967, p = .053, d = -.223, with a small effect size. These findings indicate that increasing cognitive-task variability had slightly positive influence on the writing quality for students under the PTP condition, but a detrimental effect on students' overall writing performance when students were under the OLP and CG conditions.

Two one-way ANOVA tests were performed to analyse the differences in writing quality over the three planning groups in both the simple task and the complex task. The ANOVA results showed that for the simple task, the between-group differences among the three planning types were not significant, F (2, 198) = 2.093, p = .126. Likewise, the three planning groups did not significantly differ from one another in the complex task, F (2, 198) = 1.081, p = .341. The results suggest the provision of different planning types did not significantly affect the students' writing quality when students completed either the simple task or the complex task version.

A visual representation of the effects of both cognitive-task variability and planning type can be found in Figure 4.16 above. As displayed in this figure, participants' writing quality varied differently; the overall text quality of the CG and OLP groups reduced and that of the PTP group improved, as a function of increasing cognitive-task variability.

4.3.5.4 Summary

This section summarises the effects of cognitive-task variability and planning type on students' text quality, measured by content, organisation and overall writing score (see Table 4.54).

The effects of cognitive-task variability are summarised as following. Increasing cognitive-task variability had a beneficial impact on the PTP students' writing content, writing organisation as well as the overall writing quality. Nonetheless, students under the OLP and CG conditions achieved statistically lower scores of content, organisation and overall text quality, when cognitive-task variability was raised.

Planning type had some impact in the following aspects. For the simple task, there was no significant variance between the three planning groups in terms of content, organisation and overall writing quality. With the complex task, however, the PTP group got substantially higher scores than the CG and OLP groups respectively in terms of content and organisation, but no significant differences in overall text quality were found among the three planning groups.

Significant interaction between cognitive-task variability and planning type on content, organisation and overall text quality were also identified: Students' content, organisation and writing quality on the simple versus the complex versions of the tasks were contingent respectively upon planning conditions.

Table 4.54 Summary of Effects of Cognitive-task Variability and Planning Type on Text Quality

Fluency measures	*Within-subjects differences (S vs. C)*			*Between-subjects differences (OLP vs. PTP vs. CG)*		*Interaction effects*
	OLP	*PTP*	*CG*	*Simple task*	*Complex task*	
Content	S>C	S<C	S>C	no Sig.	PTP>CG, PTP>OLP	Sig.
Organisation	S>C	S<C	S>C	no Sig.	PTP>OLP	Sig.
Overall writing quality	S>C	S<C (almost)	S>C	no Sig.	no Sig.	Sig.

Note. S = Simple task, C = Complex task, OLP = online planning group, PTP = pre-task planning group, CG = comparison group, no Sig. = no significant difference.

4.4 Chapter Summary

In this chapter, the quantitative data obtained through questionnaires, dual-task experiments and writing tasks to address RQ1 and RQ2 have been analysed and reported.

For RQ1, the results for self-ratings (i.e., participants' self-perceived task difficulty and mental effort) and for the two dual-task experiments were aligned with the intended cognitive-task variability design, that is, the designed-to be complex task had greater cognitive demands than the designed-to-be simple version.

Table 4.55 Summary of Effects of Cognitive-task Variability under Each Planning Condition

	OLP	PTP	CG
Syntactic	C/T: S>C; DC/C: S>C CP/C: S<C	CP/C: S<C	no Sig.
Lexical	WT/√2W: S>C	no Sig.	LD: S>C; WT/√2W: S>C
Accuracy	no Sig.	EP100: S>C	Err/T: S>C
Fluency	no Sig.	Fluency: S<C	no Sig.
Adequacy	Content: S>C; Organisation: S>C; Writing Quality: S>C	Content: S<C; Organisation: S<C; Writing Quality: S<C (almost)	Content: S>C; Organisation: S>C; Writing Quality: S>C

Note. S = Simple task, C = Complex task, no Sig. = no significant difference.

For RQ2, the effects of cognitive-task variability (simple versus complex tasks) on writing performance were examined under different planning conditions (see Table 4.55). For the PTP group, increasing cognitive-task variability did not influence participants' lexical complexity, but positively affected syntactic complexity, accuracy, fluency, and adequacy. For the OLP group, increases in cognitive-task variability appeared to have no effect on accuracy and fluency, but there was a decrease in lexical complexity and adequacy. A two-way influence on syntactic complexity was identified for the OLP group with reductions in C/T and DC/C and increases in CP/C, as a function of increasing cognitive-task variability. For the CG, better scores for lexical complexity and adequacy were identified in the simple task, while greater accuracy was identified in the complex task.

Table 4.56 Summary of Effects of Planning Type on Simple and Complex Tasks

	Simple task	Complex task
Syntactic	DC/C: PTP<OLP (almost)	no Sig.
Lexical	WT/√2W: PTP<OLP; PTP<CG	WT/√2W: OLP>CG
Accuracy	Err/T: PTP>CG; EP100: PTP>CG; EFC/C: PTP<CG	Err/T: PTP>CG; OLP>CG EP100: PTP>CG; OLP>CG
Fluency	no Sig.	Fluency: PTP>CG; OLP>CG
Adequacy	no Sig.	Content: PTP>CG; PTP>OLP Organisation: PTP>OLP

Note. no Sig. = no significant difference.

The different types of planning appeared to influence different aspects of EFL writing performance (see Table 4.56). For the OLP group, syntactic complexity and lexical complexity appeared to be of benefit in the simple task particularly. The CG, however, had higher scores in accuracy for both the simple and the complex tasks, while the PTP condition enabled students to achieve high scores for adequacy when the task was cognitively demanding.

The data collected for RQ3, using both quantitative and qualitative methods, are presented in the next Chapter.

Chapter 5 Relationships among Anxiety, Cognitive-task Variability, Planning Type and EFL Writing Performance

5.1 Overview

This chapter addresses the following research questions:

RQ3: *What are the relationships among students' perceived anxiety, cognitive-task variability, planning type and EFL writing production?*

(a) What is the level of second language writing anxiety reported by university students in EFL writing?

(b) What are potential effects of second language writing anxiety, cognitive-task variability and planning type on students' EFL writing?

(c) What are students' perceptions of second language writing anxiety related to cognitive-task variability?

(d) What are students' perceptions of second language writing anxiety related to planning type?

The findings from the analysis of quantitative data are reported in Section 5.2 and 5.3 to address research questions 3a and 3b. The quantitative data were collected via the two writing tasks and the SLWAI questionnaire. The results pertinent to research question 3c and 3d on the basis of the qualitative data, collected from the interview, are presented in Section 5.4.

5.2 Students' Second Language Writing Anxiety

Table 5.1 provides us with a general summary of students' L2 writing anxiety, based on their responses to the SLWAI. As shown in Table 5.1, students' total anxiety scores of the 22 items ranged from 49 to 148, with a mean of 88.12. The mean scores of the three sub-scales of SLWAI, Somatic Anxiety, Cognitive Anxiety and Avoidance Behaviour were respectively 25.56, 33.40 and 29.16.

Table 5.1 Students' Second Language Writing Anxiety

	Total	*M*	*SD*	*Minimum*	*Maximum*
Total anxiety	22	88.12	18.45	49	148
Somatic Anxiety	7	25.56	7.50	11	47
Cognitive Anxiety	8	33.40	7.62	14	53
Avoidance	7	29.16	7.00	12	49

The 201 participants in the three planning groups in phase II were classified as low, medium or high level of anxiety using the Cluster Analysis in SPSS according to students' total anxiety scores of SLWAI. The number and percentage of participants in each group are shown in Table 5.2, as well as the mean anxiety scores of the high-anxiety group (M=110.23) and the low-anxiety group (M=65.25).

Table 5.2 Cluster Analysis for High-, Medium- and Low-anxiety Groups

Anxiety group	Number	Percentage	Average anxiety
Low	52	25.9%	65.25
Medium	93	46.3%	87.60
High	56	27.9%	110.23

The high- and low-anxiety groups were identified within the three planning conditions; learners, whose anxiety score fell within the medium-anxiety group, were not included in the analysis. The information of participants in the six groups (CG-H, CG-L, PTP-H, PTP-L, OLP-H, and OLP-L) are shown in Table 5.3. Independent-samples t-tests and one-way ANOVAs were conducted to explore whether differences in anxiety among groups were statistically significant. Statistical differences in anxiety were found between the OLP-H and OLP-L groups, t (34) = 15.076, p < .001, between the PTP-H and PTP-L groups, t (34) = 12.843, p < .001, as well as between the CG-H and CG-L groups, t (34) = 15.044, p < .001. No significant differences in anxiety were identified among three high anxiety groups, F (2, 55) = .895, p = .415, or the three low anxiety groups, F (2, 51) = .301, p = .742.

Table 5.3 Students' Information in the Six Group Conditions

	N	*PT-M*	*PT-SD*	*WA-M*	*WA-SD*	*WA-t*	*df*	*WA*-Sig.
OLP-H	20	60.90	3.892	107.60	9.65	15.076	34	.000
OLP-L	16	61.75	3.173	65.50	6.25			
PTP-H	18	60.44	2.895	111.11	13.34	12.843	34	.000
PTP-L	18	62.00	3.970	64.22	7.88			
CG-H	18	61.44	3.417	112.28	10.70	15.044	34	.000
CG-L	18	62.94	4.856	66.06	7.44			

Note. PT=pre-test score, WA=writing anxiety, *t*= *t* (High-Low), *df*= *df* (High-Low), Sig. =Significance (2-tailed).

A one-way ANOVA, to examine the differences in writing proficiency among the six groups, found no statistically significant differences in pre-writing test across the six groups, F (5, 102) = .985, p = .431. The independent-samples *t*-test also identified no significant differences in the pre-test scores between participants with high anxiety and those with low anxiety, t (106) =-1.830, p=.070. These results indicate that participants in each group were similar in English writing proficiency regardless of anxiety levels.

5.3 Effects of Anxiety, Cognitive-task Variability and Planning Type

Three-way mixed ANOVAs were conducted to examine the main effects for the task (i.e., simple vs. complex), for the group (i.e., PTP, OLP, and CG) and for anxiety (i.e., high anxiety vs. low anxiety), and the interaction effects of task*group[18], task*anxiety, group*anxiety, and task*group*anxiety on each measure of students' writing production. After that, paired-samples *t*-tests were conducted for within-subjects differences (i.e., differences between simple and complex tasks across 3 planning groups and 2 levels of anxiety). A series of independent-samples *t*-tests was also administrated to detect the between-subjects differences (i.e., the comparison between high and low anxiety groups under each planning condition and between the three planning conditions under each anxiety level); these required a Bonferroni correction and the effects were reported at the significant level of .013 (.05/4=.013).

[18] *means interaction effect between task and group.

Before the statistical analyses, the assumption of normality for each of the six groups respectively under the simple and the complex tasks was first checked via skewness and kurtosis, the visual inspection of histograms, and normal Q-Q plots and box plots. Data were normally distributed, because the standardised skewness values were between 0 and +/- 3.0 and standardised kurtosis values between 0 and +/- 8.0, as recommended by Field (2009) and Kline (2011).

5.3.1 Effects on syntactic complexity

5.3.1.1 Effects on MLT

To assess the impact of language anxiety, cognitive-task variability and planning type on MLT, the assumption of normality was first checked and found to be met. Levene's *F* test for MLT in the simple task, F (5, 102) = .645, p = .666, and for MLT in the complex task, F (5, 102) = .781, p = .566, was not significant, indicating that the variance of the dependent variable in both the simple and the complex task was equal across the groups. A mixed ANOVA was, therefore, applied. Table 5.4 displays the descriptive statistics for the simple and complex tasks across six groups, with the low-anxiety groups across the CG, OLP and PTP groups achieving higher scores in MLT than the high-anxiety groups.

Table 5.4 Descriptive Statistics of Two Tasks across Six Groups for MLT

Variables	*N*	Simple task		Complex task	
		M	*SD*	*M*	*SD*
MLT-CG-H	18	11.93	2.23	11.33	2.48
MLT-CG-L	18	12.28	2.85	12.08	4.10
MLT-OLP-H	20	11.92	1.89	11.54	2.63
MLT-OLP-L	16	13.35	2.60	13.04	2.51
MLT-PTP-H	18	11.93	2.07	10.42	1.90
MLT-PTP-L	18	12.97	2.41	13.48	3.55

Note. MLT = mean length of T-unit, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H =

high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA results showed (see Table 5.5) that no interactions of task*group, task*anxiety, group*anxiety, and task*group*anxiety were found. Whereas, the main effect for anxiety, F (1, 102) = 8.536, p = .004, η_p^2 = .077, was significant, suggesting that the learners' MLT production was markedly influenced by anxiety, if the cognitive-task variability and planning type were ignored.

Table 5.5 Mixed ANOVA for Effects of Task, Anxiety and Planning on MLT

	F	p	η_p^2
task	3.563	.062	.034
group	.476	.623	.009
anxiety	8.536	.004**	.077
task*group	.041	.960	.001
task*anxiety	3.550	.062	.034
group*anxiety	.887	.415	.017
task*group*anxiety	1.863	.161	.035

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

A series of paired-samples t-tests was conducted to detect the differences between learners' MLT production on the simple task versus the complex task versions under the six conditions. The only difference to reach statistical significance was the PTP-H group, in which students produced significantly more MLT with the simple (M = 11.93, SD = 2.07) versus complex (M = 10.42, SD = 1.90) task versions, t (17) = 3.878, p = .001, d = .759; the effect size was close to large. These results indicate that increasing cognitive-task variability reduced students' MLT production for students who were highly anxious and under the PTP condition.

Independent-samples t-tests were conducted to compare the differences between the high- and low-anxiety groups under and between planning conditions. The results showed that low-anxiety students (M = 13.48, SD = 3.55) statistically outperformed their high-anxiety counterparts (*M* = 10.42, *SD* = 1.90), *t* (26.05) = -3.232, *p* = .003[19], *d* = -1.077 in terms of MLT, when they completed the complex task, but only for the PTP condition; the effect size was large. No significant differences in MLT were observed between the three planning groups at each anxiety level for either the simple or the complex tasks. These findings indicate that a low level of anxiety, in contrast to a high level of anxiety, had a positive effect on students' MLT under the PTP condition in the complex task.

5.3.1.2 Effects on C/T

The assumption of normality for C/T was evaluated and found to be satisfied. Levene's *F* test for C/T in the simple task, *F* (5, 102) = .291, *p* = .917, and in the complex task, *F* (5, 102) = 1.100, *p* = .365, showed that the assumption of homogeneity of variances was not violated for either the simple or complex task. A mixed ANOVA was, therefore, conducted. Table 5.6 shows the descriptive statistics of the simple and complex tasks across the six groups; a higher C/T score in the low-anxiety groups across the OLP, PTP and CG conditions than that in the high-anxiety groups is evident.

Table 5.6 Descriptive Statistics of Two Tasks across Six Groups for C/T

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
C/T-CG-H	18	1.42	.24	1.41	.25
C/T-CG-L	18	1.48	.29	1.53	.46
C/T-OLP-H	20	1.51	.22	1.39	.22
C/T-OLP-L	16	1.55	.22	1.46	.21
C/T-PTP-H	18	1.43	.21	1.31	.20
C/T-PTP-L	18	1.49	.22	1.48	.21

Note. C/T = clauses per T-unit, CG-H=high anxiety students under comparison group, CG-L=low anxiety students under comparison group, OLP-H=high anxiety students under online planning group, OLP-L=low anxiety students under online planning group, PTP-H=high anxiety

[19] Levene's test for MLT under PTP in the complex task was significant, *F* (1, 34) =4.555, *p*=.040, values in *Equal variances not assumed* were reported.

students under pre-task planning group, PTP-L=low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.7) found no interactions of task*group, task*anxiety, group*anxiety, and task*group*anxiety; significant main effects for task, F (1, 102) = 3.941, p = .050, η_p^2 = .037 and for anxiety, F (1, 102) = 4.043, p = .047, η_p^2 = .038, were identified. These results indicate that learners' C/T production was markedly influenced by cognitive-task variability, with planning type and anxiety ignored, and by anxiety, with cognitive-task variability and planning type ignored.

Table 5.7 Mixed ANOVA for Effects of Task, Anxiety and Planning on C/T

	F	p	η_p^2
task	3.941	.050**	.037
group	.467	.628	.009
anxiety	4.043	.047**	.038
task*group	2.311	.104	.043
task*anxiety	1.874	.174	.018
group*anxiety	.114	.892	.002
task*group*anxiety	.218	.804	.004

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples t-tests were performed and found that for students in the OLP-H group, significant differences in C/T occurred between the simple (M = 1.51, SD = .22) and complex tasks (M = 1.39, SD = .22), t (19) = 2.100, p = .049, d = .559, with a medium effect size. No significant differences between the simple and the complex tasks were found for the other five groups. These findings revealed that increasing cognitive-task variability had a negative effect on learners' syntactic complexity in terms of C/T when students were under the OLP-H condition.

Independent-samples t-tests were conducted, but no significant differences were found between the high- and low-anxiety groups under each planning condition, or between the three planning conditions under each anxiety level. These findings indicate that C/T was not significantly affected by either the anxiety level or the planning type.

5.3.1.3 Effects on DC/C

The effects of language anxiety, cognitive-task variability and planning type on DC/C were assessed and are presented in this section. A mixed ANOVA was computed, since the assumption of normality was met, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* test for the simple task, F (5, 102) = .365, p = .872, and for the complex task, F (5, 102) = .215, p = .955). The descriptive statistics of the simple and complex tasks across the six groups are presented in Table 5.8. Students' DC/C performance was the lowest in the CG-H group for the simple task and in the PTP-H group for the complex task, and the highest in the OLP-L group for both the simple and complex tasks.

Table 5.8 Descriptive Statistics of Two Tasks across Six Groups for DC/C

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
DC/C-CG-H	18	.28	.10	.26	.11
DC/C-CG-L	18	.30	.11	.30	.14
DC/C-OLP-H	20	.32	.09	.27	.10
DC/C-OLP-L	16	.36	.09	.32	.10
DC/C-PTP-H	18	.29	.09	.22	.11
DC/C-PTP-L	18	.32	.09	.31	.11

Note. DC/C = dependent clauses per clause, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

According to the mixed ANOVA (see Table 5.9), whereas no interactions of task*group, task*anxiety, group*anxiety, and task*group*anxiety were found, significant main effects for task, F (1, 102) = 8.342, p = .005, η_p^2 = .076 and for anxiety, F (1, 102) = 6.143, p = .015, η_p^2 = .057, were identified. These results indicate that participants' DC/C production was markedly influenced by cognitive-task variability, with planning type and anxiety ignored, and by anxiety, with cognitive-task variability and planning type ignored.

Table 5.9 Mixed ANOVA for Effects of Task, Anxiety and Planning on DC/C

	F	Sig.	η_p^2
task	8.342	.005**	.076
group	1.633	.200	.031
anxiety	6.143	.015**	.057
task*group	1.395	.253	.027
task*anxiety	1.728	.192	.017
group*anxiety	.262	.770	.005
task*group*anxiety	.348	.707	.007

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples *t*-tests were conducted to examine the location of the within-subjects differences. The results showed that students under the OLP-H condition produced statistically more DC/C in the simple task (*M* = .32, *SD* = .09) than in the complex task (*M* = .27, *SD* = .10), *t* (19) = 2.114, *p* =.048, *d* = .556, and the effect size was medium. Likewise, significant differences in DC/C were found for the PTP-H group: simple task (*M* = .29, *SD* = .09) versus complex task (*M* = .22, *SD* = .11), *t* (17) = 2.418, *p* = .027, *d* = .760, with an effect size close to large. No significant differences were found between the simple and complex tasks for other four groups. These findings suggest that increasing cognitive-task variability had negative effects on DC/C, for students with high anxiety under the OLP and PTP conditions.

Independent-samples *t*-tests were conducted, but no statistical significances were found between the high- and low-anxiety groups under each planning condition, or between the three planning conditions under each anxiety level. These findings indicate that DC/C was not significantly affected by either the anxiety level or the planning type.

5.3.1.4 Effects on T/S

Table 5.10 below illustrates the descriptive statistics of the simple and complex tasks across the six groups. The assumption of normality was met, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* test for the simple task, *F* (5, 102) = 1.102, *p* = .364, and for the complex task, *F* (5, 102) = .269, *p* = .929). A mixed ANOVA was, therefore, conducted.

Table 5.10 Descriptive Statistics of Two Tasks across Six Groups for T/S

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
T/S-CG-H	18	1.14	.12	1.13	.12
T/S-CG-L	18	1.10	.10	1.14	.16
T/S-OLP-H	20	1.10	.08	1.12	.10
T/S-OLP-L	16	1.10	.10	1.15	.12
T/S-PTP-H	18	1.12	.10	1.12	.10
T/S-PTP-L	18	1.08	.08	1.12	.13

Note. T/S = T-units per sentence, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

As shown in Table 5.11, neither main effects for task, group, or anxiety, nor interaction effects of task*group, task*anxiety, group*anxiety, or task*group*anxiety were identified. These findings suggest that participants' T/S performance did not differ significantly as a function of cognitive-task variability, planning type and anxiety level.

Table 5.11 Mixed ANOVA for Effects of Task, Anxiety and Planning on T/S

	F	Sig.	η_p^2
task	2.867	.093	.027
group	.602	.549	.012
anxiety	.147	.703	.001
task*group	.252	.778	.005
task*anxiety	1.654	.201	.016
group*anxiety	.507	.604	.010
task*group*anxiety	.107	.899	.002

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

5.3.1.5 Effects on MLC

Table 5.12 depicts the descriptive statistics of the simple and complex tasks across six groups. Low-anxiety participants under the OLP and PTP conditions produced higher scores in MLC respectively than their high-anxiety counterparts for both the simple and complex tasks, while the CG-H group had higher scores than the CG-L group for both the simple and complex tasks.

Table 5.12 Descriptive Statistics of Two Tasks across Six Groups for MLC

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
MLC-CG-H	18	8.43	.92	8.02	1.02
MLC-CG-L	18	8.30	.81	7.90	.79
MLC-OLP-H	20	7.93	.85	8.33	1.43
MLC-OLP-L	16	8.60	.97	8.90	1.11
MLC-PTP-H	18	8.39	1.37	7.93	.72
MLC-PTP-L	18	8.78	1.43	9.08	1.73

Note. MLC = mean length of clause, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H =

high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

To investigate the impact of language anxiety, cognitive-task variability and planning type on MLC, the assumption of normality was evaluated and found to be met. Levene's *F* test for MLC in the simple task, F (5, 102) = .971, p = .440, and for MLC in the complex task, F (5, 102) = 1.530, p = .187, was not significant, indicating that the variance of the dependent variable in both the simple and the complex task was equal across the groups. A mixed ANOVA was, hence, applied.

Table 5.13 Mixed ANOVA for Effects of Task, Anxiety and Planning on MLC

	F	Sig.	η_p^2
task	.167	.684	.002
group	1.492	.230	.028
anxiety	5.021	.027**	.047
task*group	3.569	.032**	.065
task*anxiety	.909	.343	.009
group*anxiety	2.210	.115	.042
task*group*anxiety	1.334	.268	.025

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

The mixed ANOVA results identified (Table 5.13) that both the main effect for anxiety, F (1, 102) = 5.021, p = .027, η_p^2 = .047 and the interaction for task*group, F (2, 102) = 3.569, p = .032, η_p^2 = .065 were significant. Nevertheless, neither the main effects for task, and for group, nor the interaction of task*anxiety, group*anxiety and task*group*anxiety were significant. The results suggest that participants' MLC performance was influenced markedly by anxiety, if the cognitive-task variability and planning type were ignored. Additionally, the MLC performance of participants in the simple and complex task versions differed across the CG, PTP and OLP groups.

Paired-samples *t*-tests were conducted, but no significant differences in MLC were detected between the simple and complex tasks under the six conditions. These results indicate that cognitive-task variability rarely had an effect on students' MLC performance.

A series of independent-samples *t*-tests was applied to compare the differences between the high- and low-anxiety groups under and between planning conditions. With the complex task, significant differences in MLC were found between the OLP-L (M = 8.90, SD = 1.11) and CG-L groups (M = 7.90, SD = .79), $t(32)$= 3.042, p = .005, d = -1.045, and between the PTP-L (M = 9.08, SD = 1.73) and CG-L groups (M = 7.90, SD = .79), t (34) = 2.634, p = .013, d = -.878, for both the effect sizes were large. No significant differences were detected, however, among CG-L, OLP-L and PTP-L, when students completed the simple task, or among the CG-H, OLP-H, and PTP-H groups for either the simple or the complex tasks. No significant differences were identified in the comparisons of MLC between the high- and low-anxiety levels under the OLP, PTP and CG conditions. These results indicate that for the low-anxiety students, providing online planning and pre-task planning had beneficial impacts on their MLC production in the complex task. For highly anxious students, no significant effects on students' MLC performance were, however, found across the CG, OLP, and PTP groups in either simple or complex tasks. The high- and low-anxiety participants' MLC was similar under the CG, OLP and PTP conditions.

5.3.1.6 Effects on CP/C

Table 5.14 presents the descriptive statistics of the simple and complex tasks across the six groups. The assumption of normality for CP/C was evaluated and found to be satisfied. Levene's F test for CP/C in the simple task, F (5, 102) = .521, p = .760, and in the complex task, F (5, 102) = 1.405, p = .229, was not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA was, therefore, conducted.

Table 5.14 Descriptive Statistics of Two Tasks across Six Groups for CP/C

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
CP/C-CG-H	18	.34	.15	.35	.11
CP/C-CG-L	18	.35	.13	.29	.09
CP/C-OLP-H	20	.29	.10	.36	.16
CP/C-OLP-L	16	.34	.11	.42	.12
CP/C-PTP-H	18	.32	.13	.39	.10
CP/C-PTP-L	18	.35	.13	.35	.11

Note. CP/C = coordinate phrases per clause, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.15) did not identify either the main effects for group and for anxiety, or the interaction of task*anxiety, group*anxiety and task*group*anxiety, whereas, a main effect for task, F (1, 102) = 3.954, p = .049, η_p^2 = .037, and the interaction between task and group, F (2, 102) = 3.413, p = .037, η_p^2 = .063 were identified as significant. The results indicate that participants' CP/C performance was significantly affected by cognitive-task variability, if the planning type and anxiety were ignored. Additionally, participants' CP/C performance in the simple and complex tasks differed across the CG, PTP and OLP groups.

Table 5.15 Mixed ANOVA for Effects of Task, Anxiety and Planning on CP/C

	F	Sig.	η_p^2
task	3.954	.049**	.037
group	.712	.493	.014
anxiety	.177	.675	.002
task*group	3.413	.037**	.063
task*anxiety	2.082	.152	.020
group*anxiety	2.254	.110	.042
task*group*anxiety	.825	.441	.016

Note. Task = cognitive-task variability, group = planning type.
***p*<.05; ****p*<.001

Paired-samples *t*-tests were performed and it was found that for students under the OLP-L condition, CP/C produced in the complex task (*M* = .42, *SD* = .12) substantially outnumbered those performed in the simple task (*M* = .34, *SD* = .11), *t* (15) = -2.162, *p* = .047, *d* = -.703, with an effect size close to large. No significant differences between the simple and complex tasks were found for the other five groups. These findings suggest that increasing cognitive-task variability benefited students' CP/C performance under the OLP-L condition.

Independent-samples *t*-tests showed that for the complex task, but not the simple task, the OLP-L group (*M* = .42, *SD* = .12) significantly outperformed the CG-L group (*M* = .29, *SD* = .09) in CP/C, *t* (32) = 3.710, *p* = .001, *d* = 1.275 and the effect size was large. No significant differences were found between the three planning conditions for the high-anxiety level, or between the high- and low-anxiety groups under each planning condition. These results suggest that for the low-anxiety students, online planning benefited CP/C production in the complex task, while for the high-anxiety students, there were no significant effects on students' CP/C performance across the CG, OLP and PTP groups. The high- and low-anxiety participants, under the CG, OLP and PTP conditions, made similar performance in CP/C.

5.3.1.7 Effects on CN/C

The effects of language anxiety, cognitive-task variability and planning type on CN/C were assessed and are presented in this section, with the descriptive statistics of the simple and complex tasks across the six groups displayed in Table 5.16. Students' CN/C performance was the lowest in the PTP-H group in the complex task, and the highest in the PLP-L group in the simple task.

Table 5.16 Descriptive Statistics of Two Tasks across Six Groups for CN/C

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
CN/C-CG-H	18	.86	.20	.79	.26
CN/C-CG-L	18	.82	.21	.83	.23
CN/C-OLP-H	20	.82	.15	.84	.25
CN/C-OLP-L	16	.93	.26	.93	.24
CN/C-PTP-H	18	.85	.21	.73	.19
CN/C-PTP-L	18	.96	.24	.95	.33

Note. CN/C = complex nominals per clause, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

A mixed ANOVA was computed, since the assumption of normality was met, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* test for the simple task, F (5, 102) = .584, p = .712, or the complex task, F (5, 102) = .856, p = .513). The mixed ANOVA (see Table 5.17) did not identify either interactions for task*group, task*anxiety, group*anxiety, and task*group*anxiety, or the main effects for task and for group. A main effect for anxiety, F (1, 102) = 5.496, p = .021, η_p^2 = .051, was identified, suggesting that anxiety significantly affected students' CN/C production, if the cognitive-task variability and planning type were ignored.

Table 5.17 Mixed ANOVA for Effects of Task, Anxiety and Planning on CN/C

	F	Sig.	η_p^2
task	1.404	.239	.014
group	.942	.393	.018
anxiety	5.496	.021**	.051
task*group	.784	.459	.015
task*anxiety	1.397	.240	.014
group*anxiety	1.493	.230	.028
task*group*anxiety	.822	.442	.016

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples *t*-tests were conducted to examine where the within-subjects differences were located, and found that students under the PTP-H condition produced statistically more CN/C in the simple task (M = .85, SD = .21) than in the complex task (M = .73, SD = .19), t (17) = 2.333, p = .032, d = .731, with an effect size that was nearly large. No significant differences were found for the other five groups between the simple and complex tasks. These findings suggest that increasing cognitive-task variability had a negative effect on CN/C production for students with high anxiety under the PTP condition.

Independent-samples *t*-tests did not identify any statistical significances between the high- and low-anxiety groups under each planning condition, or between the three planning conditions under each anxiety level. These findings suggest that CN/C was not significantly affected by the anxiety level and planning type.

5.3.1.8 Summary

In this section, the effects of language anxiety, cognitive-task variability and planning type on syntactic complexity are summarised.

Neither interactions for task*group*anxiety, nor interactions between anxiety and task, and between anxiety and groups were significant across the seven measures of syntactic complexity, suggesting that participants' syntactic performance in the simple and complex task versions and across the CG, PTP and OLP groups were not affected by either high- or low-anxiety levels. Interaction effects of task*group, and main effects for anxiety and for tasks were found significant in some measures of syntactic complexity.

The within-subjects differences across the three planning conditions and two anxiety groups are summarised first. Table 5.18 shows that for high-anxiety students, increasing cognitive-task variability had a detrimental effect on C/T and DC/C under the OLP condition, and on MLT, DC/C and CN/C under the PTP condition. The effect sizes under the PTP group were larger than those under the OLP group. By contrast, the OLP-L group' syntactic-complexity performance in terms of CP/C was better in the complex task than the simple task, and the effect size was close to large.

Table 5.18 Summary of Significant Within-subjects Differences for Syntactic Complexity

Anxiety group	Planning	Measures	Differences	Sig.	*d*
High anxiety	OLP	C/T	Simple > Complex	.049	.559
		DC/C	Simple > Complex	.048	.556
	PTP	MLT	Simple > Complex	.001	.759
		DC/C	Simple > Complex	.027	.760
		CN/C	Simple > Complex	.032	.731
Low anxiety	OLP	CP/C	Simple < Complex	.047	-.703

Table 5.19 Summary of Significant Between-subjects Differences for Syntactic Complexity

Task	Measures	Differences	Sig.	*d*
Simple	/	/	/	/
Complex	MLT	PTP-H < PTP-L	.003	-1.077
	MLC	OLP-L > CG-L	.005	1.045
		PTP-L > CG-L	.013	.878
	CP/C	OLP-L > CG-L	.001	1.275

Significant between-subjects differences were found when students completed the complex task compared with the simple task. Specifically, a low level of anxiety appeared to help the PTP group produce more MLT in the complex task than those with a high level of anxiety. Students in the planning groups (i.e., OLP and PTP), with a low-anxiety level, significantly outperformed the CG-L group in MLC. Likewise, the CP/C of the OLP-L group was better performance than the CG-L group, for the complex task.

5.3.2 Effects on lexical complexity

The findings of the effects of language anxiety, cognitive-task variability and planning type on students' lexical complexity are reported in this section. Lexical complexity was measured by lexical density (LD) and lexical diversity (WT/√2W).

5.3.2.1 Effects on LD

Table 5.20 illustrates the descriptive statistics of the simple and complex tasks across the six groups. The data on LD were analysed with the mixed ANOVA test, since the assumption of normality was tenable, and the assumption of homogeneity of variances was not violated for either the simple or complex task (Levene's *F* tests for LD in the simple task, F (5, 102) = .095, p = .993, and for LD in the complex task, F (5, 102) = .842, p = .523).

Table 5.20 Descriptive Statistics of Two Tasks across Six Groups for LD

Variables	N	Simple task		Complex task	
		M	SD	M	SD
LD-CG-H	18	.518	.03	.504	.03
LD-CG-L	18	.517	.04	.506	.03
LD-OLP-H	20	.506	.03	.508	.03
LD-OLP-L	16	.513	.03	.523	.04
LD-PTP-H	18	.503	.03	.512	.04
LD-PTP-L	18	.506	.04	.508	.03

Note. CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.21) did not identify either interactions of task*anxiety, group*anxiety, and task*group*anxiety, or the main effects for task, group, and anxiety for LD. An almost significant interaction effect for task*group, F (1, 102) = 3.045, p = .052, $\eta_p{}^2$ = .056, was found, suggesting that participants' LD production in the simple and complex task versions differed across the PTP, OLP and CG conditions.

Table 5.21 Mixed ANOVA for Effects of Task, Anxiety and Planning on LD

	F	Sig.	η_p^2
task	.004	.950	.000
group	.395	.674	.008
anxiety	.528	.469	.005
task*group	3.045	**.052**	.056
task*anxiety	.049	.826	.000
group*anxiety	.608	.546	.012
task*group*anxiety	.342	.711	.007

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples *t*-tests, to compare the within-subjects differences across the six groups, identified statistically significant differences between the simple (*M* = .518, *SD* = .03) and the complex tasks, (*M* = .504, *SD* = .03), *t* (17) = 2.898, *p* = .010, *d* = .496 in LD for the CG-H group. The effect size was small. No significant differences between the simple and complex tasks were found for the other five groups. These findings suggest that, for the students in the CG-H group, increasing cognitive-task variability had a negative effect on LD production.

Independent-samples *t*-tests were conducted, but no significant differences in LD were found between the high- and low-anxiety groups under each planning condition, or between the three planning conditions under each anxiety level. These results suggest that LD was not significantly affected by either the anxiety level and or the planning type.

5.3.2.2 *Effects on WT/ √2W*

The descriptive statistics of the simple and complex tasks across the six groups, presented in Table 5.22, show that students with low anxiety produced better performance in WT/√2W than those with high anxiety across the OLP, PTP and CG conditions.

Table 5.22 Descriptive Statistics of Two Tasks across Six Groups for WT/√2W

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
WT/√2W-CG-H	18	5.47	.36	5.04	.28
WT/√2W-CG-L	18	5.50	.49	5.29	.66
WT/√2W-OLP-H	20	5.45	.43	5.21	.51
WT/√2W-OLP-L	16	5.63	.62	5.48	.49
WT/√2W-PTP-H	18	4.89	.25	4.95	.31
WT/√2W-PTP-L	18	5.49	.44	5.57	.58

Note. CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The assumption of normality was evaluated and found to be met. The homogeneity of variances assumption was, however, violated, as Levene's *F* test for WT/√2W in the simple task, F (5, 102) = 2.528, p = .034, and for WT/√2W in the complex task, F (5, 102) = 3.642, *p* = .005, was significant. A mixed ANOVA was, therefore, not applied. Instead, the paired-samples *t*-tests and independent-samples *t*-tests were conducted to explore the impact of language anxiety, cognitive-task variability and planning type on WT/√2W.

Paired-samples *t*-tests were conducted to detect the differences in WT/√2W between the simple and complex tasks under the six conditions. According to the findings, significant differences in WT/√2W were found in the OLP-H group between the simple (*M* = 5.45, *SD* = .43) and the complex tasks (*M* = 5.21, *SD* = .51), *t* (19) = 2.157, *p* = .044, *d* = .506, with a medium effect size. The same pattern was found in the CG-H group; WT/√2W produced in the simple task (*M* = 5.47, *SD* = .36) statistically outnumbered those produced in the complex task (*M* = 5.04, *SD* = .28), *t* (17) = 4.592, *p* < .001, *d* = 1.353, with a large effect size. No significant differences in WT/√2W were identified for the other four groups between the simple and complex tasks. These findings indicate that increasing cognitive-task variability negatively influenced students' WT/√2W production under the OLP-H and CG-H conditions.

A series of independent-samples *t*-tests was applied to compare the differences between the high- and low-anxiety groups under and between planning conditions. The results are shown in Table 5.23.

Table 5.23 Summary of Statistically Significant Group Differences from the Independent *t*-tests for WT/√2W

task	Group(I)----(J)	MD.(I-	Std.Er	*t*	Sig.	*d*
Simple	PTP-H - PTP-L	-.60	.12	-5.022	.000***	-1.674
Complex	PTP-H - PTP-L	-.61	.15	-3.966	.001[20]**	-1.322
Simple	CG-H - PTP-H	.59	.10	5.683	.000***	1.894
Simple	OLP-H - PTP-H	.56	.11	4.860	.000***	1.579

Note. MD. = Mean differences, Std = Standard, Er. = Error, H=high anxiety, L=low anxiety, LD = lexical density.
**p<.013; ***p<.001

As shown in Table 5.23, for the simple task, PTP-L students (M = 5.49, SD = .44) outperformed their PTP-H counterparts (M = 4.89, SD = .25) in terms of WT/√2W, t (34) = -5.022, p < .001, d = -1.674, with a large effect size. The same pattern was found for the complex task, with significant differences in WT/√2W identified between the PTP-H (M = 4.95, SD = .31) and PTP-L (M = 5.57, SD = .58) groups, t (25.78) = -3.966, p = .001, d = -1.322; the effect size was large. These findings reveal that a low level of anxiety enhanced PTP students' production of a higher level of WT/√2W in both the simple and complex tasks compared with high-anxiety level students.

When completing the simple task, the CG-H group (M = 5.47, SD = .36) significantly outperformed the PTP-H group (M = 4.89, SD = .25), p < .001, d = 1.894 in terms of WT/√2W production, with a large effect size. The OLP-H group (M = 5.45, SD = .43), similarly, produced more WT/√2W than the PTP-H group (M = 4.89, SD = .25) in the simple task, p <.001, d = 1.579, with a large effect size. These findings suggest that the CG-H and OLP-H groups both performed better with WT/√2W than the PTP-H group, when completing the simple task.

[20] Levene's test for WT/√2W in the complex task was significant, F (1, 34) =7.217, p=.011, values in *Equal variances not assumed* were reported.

5.3.2.3 Summary

In this section, the effects of language anxiety, cognitive-task variability and planning type on lexical complexity are summarised.

Table 5.24 Summary of Significant Within-subjects Differences for Lexical Complexity

Anxiety	Planning	Measures	Differences	Sig.	*d*
High anxiety	OLP	WT/√2W	Simple > Complex	.044	.506
	CG	LD	Simple > Complex	.010	.496
		WT/√2W	Simple > Complex	.000	1.353

The within-subjects differences across the three planning conditions and the two anxiety groups are summarised first (see Table 5.24). Increasing cognitive-task variability had a detrimental effect on WT/√2W under the OLP-H condition and on LD and WT/√2W under the CG-H condition. The effect size for WT/√2W in the CG-H group was larger than that in the OLP-H group. By contrast, for low-anxiety group, cognitive-task variability had little effect on students' lexical complexity performance across the OLP, PTP and CG conditions.

Table 5.25 Summary of Significant Between-subjects Differences for Lexical Complexity

Task	Measures	Differences	Sig.	*d*
Simple	WT/√2W	PTP-H <PTP-L	.000	-1.674
	WT/√2W	CG-H > PTP-H	.000	1.894
	WT/√2W	OLP-H > PTP-H	.000	1.579
Complex	WT/√2W	PTP-H < PTP-L	.001	-1.322

The statistical significance between the high- and low-anxiety groups under and between planning conditions is summarised in Table 5.25. A low level of anxiety appeared to benefit the PTP group with WT/√2W performance for both the simple and complex task. Both the CG-H and OLP-H groups achieved significantly higher scores for WT/√2W than the PTP-H group, when completing the simple task.

5.3.3 Effects on accuracy

Students' accuracy performance was measured by Err/T, EP100 and EFC/C. This section reports the effects of language anxiety, cognitive-task variability and planning type on students' EFL accuracy.

5.3.3.1 Effects on Err/T

Table 5.26 depicts the descriptive statistics of the simple and complex tasks across the six groups. The assumption of normality for Err/T was evaluated and found to be satisfied. Levene's *F* test for Err/T in the simple task, F (5, 102) = .629, p = .678, and in the complex task, F (5, 102) = 1.987, p = .087, showed that the homogeneity of variances assumption was not violated for either the simple or complex task. A mixed ANOVA was, therefore, conducted.

Table 5.26 Descriptive Statistics of Two Tasks across Six Groups for Err/T

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
Err/T-CG-H	18	.57	.24	.58	.28
Err/T-CG-L	18	.67	.31	.52	.23
Err/T-OLP-H	20	.68	.26	.69	.33
Err/T-OLP-L	16	.81	.28	.87	.48
Err/T-PTP-H	18	.81	.27	.58	.20
Err/T-PTP-L	18	.81	.28	.74	.42

Note. CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.27) identified significant differences in Err/T for main effect for task, F (1, 102) = 5.219, p = .024, η_p^2 = .049 and for group, F (2, 102) = 4.799, p = .010, η_p^2 = .086, but not for anxiety. Whereas the interaction between task and group, F (2, 102) = 3.591, p = .031, η_p^2 = .066 was significant, there were no significant interactions for task*anxiety, group*anxiety and task*group*anxiety. These results suggest that students' Err/T differences between the simple and the complex versions were dependent on planning type but not anxiety. Students' Err/T was markedly influenced by cognitive-task variability, with planning type and anxiety ignored, and by group, with cognitive-task variability and anxiety ignored.

Table 5.27 Mixed ANOVA for Effects of Task, Anxiety and Planning on Err/T

	F	Sig.	η_p^2
task	5.219	.024**	.049
group	4.799	.010**	.086
anxiety	2.448	.121	.023
task*group	3.591	.031**	.066
task*anxiety	.107	.744	.001
group*anxiety	.590	.556	.011
task*group*anxiety	2.780	.067	.052

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples *t*-tests were performed and found that PTP-H students made more Err/T in the simple task (M = .81, SD = .27) than the complex task (M = .58, SD = .20), t (17) 4.221, p = .001, d = .944. The effect size was large. The same pattern was identified for the CG-L group, where more Err/T was made in the simple task (M = .67, SD = .31) than the complex task (M = .52, SD = .23), t (17) = 2.201, p = .042, d = .537. The effect size was medium. No significant differences in Err/T between the simple and complex tasks were found in other four groups. These findings suggest that the complex task led to students under the PTP-H and CG-L conditions make fewer Err/T than the simple task.

Independent-samples *t*-tests showed that for the simple task, the PTP-H group made significantly more Err/T than the CG-H group, t (34) = 2.860, p = .007, d = .953 and the effect size was large. For the complex task, the OLP-L group (M = .87, SD = .48) produced more Err/T, t (32) =2.779, p = .009, d = .955 than the CG-L group (M = .52, SD = .23), with a large effect size. No significant differences were found between the high- and low-anxiety groups under each planning condition. These results indicate that compared to CG-H students, PTP-H students made significantly more Err/T with the simple task. When the task was cognitively demanding, CG-L students' accuracy was better in Err/T than OLP-L students.

5.3.3.2 Effects on EP100

Table 5.28 presents the descriptive statistics of the simple and complex tasks across the six groups. According to the table, students' EP100 performance across the six groups for the simple and complex tasks did not differ substantially. The data on EP100 were analysed with the mixed ANOVA test, because the assumption of normality was tenable, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's F tests for EP100 in the simple task, F (5, 102) = 1.626, p = .160, and for EP100 in the complex task, F (5, 102) = .409, p = .842).

Table 5.28 Descriptive Statistics of Two Tasks across Six Groups for EP100

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
EP100-CG-H	18	.05	.02	.05	.02
EP100-CG-L	18	.06	.03	.04	.02
EP100-OLP-H	20	.06	.02	.06	.03
EP100-OLP-L	16	.06	.02	.07	.03
EP100-PTP-H	18	.07	.03	.06	.02
EP100-PTP-L	18	.06	.02	.05	.03

Note. CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.29) identified no interactions for task*anxiety, group*anxiety, or task*group*anxiety for EP100, whereas, a significant interaction effect for task*group, F (1, 102) = 3.783, p = .026, η_p^2 = .069, was found, suggesting that participants' EP100 in the simple and complex task versions differed across the CG, PTP and OLP groups. Only a main effect for group was found, F (2, 102) = 3.753, p = .027, η_p^2 = .069, indicating that planning type substantially affected students' EP100 performance, if cognitive-task variability and anxiety were not considered.

Table 5.29 Mixed ANOVA for Effects of Task, Anxiety and Planning on EP100

	F	Sig.	η_p^2
task	2.871	.093	.027
group	3.753	.027**	.069
anxiety	.002	.963	.000
task*group	3.783	.026**	.069
task*anxiety	.581	.448	.006
group*anxiety	.523	.594	.010
task*group*anxiety	1.794	.172	.034

Note. Task = cognitive-task variability, group = planning type.
$^{**}p<.05$; $^{***}p<.001$

Paired-samples *t*-tests, performed to compare the within-subjects differences across the six groups, showed that PTP-H students made significantly more EP100 in the simple task (*M* = .07, *SD* = .03) than in the complex task (*M* = .06, *SD* = .02), *t* (17) = 2.380, *p* = .029, *d* = .498, with a medium effect size. No significant differences between the simple and complex tasks were found in the other five groups. These findings suggest that an increase in cognitive-task variability benefited high-anxiety students' accuracy in EP100 under the PTP condition.

Independent-samples *t*-tests were conducted, but no significant differences in EP100 were found between the high- and low-anxiety groups under each planning condition. Nonetheless, when students completed the simple task, EP100 made by the PTP-H group (*M* = .07, *SD* = .03) significantly outnumbered those produced by the CG-H group (*M* = .05, *SD* = .02), *t* (34) = 2.792, *p* = .009, *d* = -.931, with a large effect size. For the complex task, the OLP-L group (*M* = .07, *SD* = .03) produced more EP100, *t* (32) = 2.664, *p* = .012, *d* = .915 than the CG-L group (*M* = .04, *SD* = .02). The effect size was large. These findings suggest that PTP-H students made significantly more EP100 than CG-H students, when completing the simple task, while for the complex task, the CG-L group got better accuracy performance than the OLP-L group in terms of EP100.

5.3.3.3 Effects on EFC/C

To assess the impacts of language anxiety, cognitive-task variability and planning type on EFC/C, the assumption of normality was evaluated and found to be met. Levene's *F* test for EFC/C in the simple task, F (5, 102) = 1.704, p = .140, and for EFC/C in the complex task, F (5, 102) = 2.214, p = .058, was not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA was, hence, applied.

Table 5.30 shows the descriptive statistics of the simple and complex tasks across the six groups.

Table 5.30 Descriptive Statistics of Two Tasks across Six Groups for EFC/C

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
EFC/C-CG-H	18	.68	.10	.68	.11
EFC/C-CG-L	18	.65	.13	.71	.10
EFC/C-OLP-H	20	.64	.08	.62	.16
EFC/C-OLP-L	16	.61	.12	.59	.17
EFC/C-PTP-H	18	.56	.12	.62	.13
EFC/C-PTP-L	18	.63	.19	.68	.28

Note. CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA results in Table 5.31 showed that neither the main effects for task, group as well as anxiety, nor the interactions for task*group, task*anxiety, group*anxiety, and task*group*anxiety were significant. The non-significant interaction effects suggest that participants' EFC/C in the simple versus the complex versions was not dependent on planning type and anxiety level.

Table 5.31 Mixed ANOVA for Effects of Task, Anxiety and Planning on EFC/C

	F	Sig.	η_p^2
task	1.722	.192	.017
group	2.699	.072	.050
anxiety	.127	.723	.001
task*group	2.030	.137	.038
task*anxiety	.216	.643	.002
group*anxiety	1.494	.229	.028
task*group*anxiety	.562	.572	.011

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples t-tests revealed that PTP-H students made more EFC/C in the complex task (M = .62, SD = .13) than in the simple task (M = .56, SD = .12), t (17) = -2.133, p = .048, d = -.503, with a medium effect size. No significant differences were detected between the simple and complex tasks for the other five groups. These findings suggest that increases in cognitive-task variability had a beneficial influence on accuracy in terms of EFC/C for high-anxiety students under the PTP condition.

A series of independent-samples t-tests was applied to compare the differences between the high- and low-anxiety groups under and between planning conditions. The results showed that when students completed the simple task, significant differences in EFC/C were found between the CG-H (M = .68, SD = .10), and PTP-H groups (M = .56, SD = .12), t (34) = 3.325, p = .002, d = 1.108. The effect size was large. These results indicate that in the simple task, the accuracy of the CG-H group was better, as measured by EFC/C, than the PTP-H group.

5.3.3.4 Summary

In this section, the effects of language anxiety, cognitive-task variability and planning type on accuracy are summarised.

No significant interactions for task*group*anxiety, group*anxiety, task* anxiety were identified across the three measures of accuracy, suggesting that participants' accuracy in the simple and complex tasks and across the CG, PTP and OLP groups did not significantly differ between high- and low-anxiety students.

Table 5.32 Summary of Significant Within-subjects Differences for Accuracy

Anxiety	Planning	Measures	Differences	Sig.	*d*
High anxiety	PTP	Err/T	Simple > Complex	.001	.944
		EP100	Simple > Complex	.029	-.498
		EFC/C	Simple < Complex	.048	-.503
Low anxiety	CG	Err/T	Simple > Complex	.042	.537

The within-subjects differences across the three planning conditions and two anxiety groups are summarised in Table 5.32. Increases in cognitive-task variability had a positive effect on PTP-H students' accuracy performance in terms of Err/T, EP100 and EFC/C. For the low-anxiety group, increasing cognitive-task variability under the CG condition also led to fewer Err/T.

Table 5.33 Summary of Significant Between-subjects Differences for Accuracy

Task	Measures	Differences	Sig.	*d*
Simple	Err/T	PTP-H > CG-H	.007	.953
	EP100	PTP-H > CG-H	.009	.931
	EFC/C	PTP-H < CG-H	.002	1.108
Complex	Err/T	OLP-L > CG-L	.009	.955
	EP100	OLP-L > CG-L	.012	.915

The statistical significances between the high- and low-anxiety groups under and between planning conditions are summarised in Table 5.33. With the simple task, the CG-H group had better accuracy than the PTP-H group in terms of Err/T, EP100 and EFC/C. Similarly, students in the CG-L group performed better in accuracy, in terms of Err/T and EP100, than those in the OLP-L group, when completing the complex task.

5.3.4 Effects on fluency

The findings for the effects of language anxiety, cognitive-task variability and planning type on students' fluency are reported in this section. The descriptive statistics of students' writing fluency across the six groups for both simple and complex tasks show that low-anxiety students wrote more fluently than the high-anxiety students in both the simple and complex task versions, and under the three planning conditions (Table 5.34).

Table 5.34 Descriptive Statistics of Two Tasks across Six Groups for Fluency

Variables	N	Simple task		Complex task	
		M	SD	M	SD
FL-CG-H	18	5.26	.75	5.38	.65
FL-CG-L	18	5.55	1.10	5.56	1.00
FL-OLP-H	20	5.77	.95	5.61	1.10
FL-OLP-L	16	6.02	1.03	5.76	.83
FL-PTP-H	18	5.25	.85	5.77	.82
FL-PTP-L	18	5.85	1.38	6.53	1.28

Note. FL = fluency, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The assumption of normality was evaluated and found to be met. Levene's F test for writing fluency in the simple task, F (5, 102) = .783, p = .564, and in the complex task, F (5, 102) = 1.534, p = .186, was not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA, therefore, was applied to assess the impacts of language anxiety, cognitive-task variability and planning type on students' fluency.

Table 5.35 Mixed ANOVA for Effects of Task, Anxiety and Planning on Fluency

	F	Sig.	η_p^2
task	2.981	.087	.028
group	2.282	.107	.043
anxiety	4.659	.033**	.044
task*group	7.143	.001**	.123
task*anxiety	.009	.923	.000
group*anxiety	.824	.441	.016
task*group*anxiety	.282	.755	.005

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

The mixed ANOVA (see Table 5.35) found significant main effects for anxiety, F (1, 102) = 4.659, p = .033, η_p^2 = .044, but not for task and group. The interaction between task and group, F (2, 102) = 7.143, p = .001, η_p^2 = .123, was significant. The interactions for task* anxiety, group*anxiety and task*group*anxiety, however, were not significant. The results indicate that participants' fluency was influenced markedly by anxiety, if cognitive-task variability and planning type were ignored. Fluency in the simple and complex task versions differed across the PTP, OLP and CG conditions.

Paired-samples t-tests detected differences in fluency between the simple and complex tasks under the six conditions. Significant differences in fluency were found in the PTP-H group between the simple (M = 5.25, SD = .85) and complex tasks (M = 5.77, SD = .82), t (17) = -3.550, p = .002, d= -.616, associated with a medium effect size. Similarly, students under the PTP-L condition wrote significantly more fluently in the complex task (M = 6.53, SD = 1.28) than in the simple task (M = 5.85, SD = 1.38), t (17) = -3.465, p = .003, d= -.513. The effect size was medium. No significant differences were detected for participants' fluency production on the simple task versus the complex task versions under the OLP and CG planning conditions. The findings indicate that increasing cognitive-task variability positively influenced fluency for the PTP-H and PTP-L groups.

A series of independent-samples t-tests was applied to compare the between-subjects differences in fluency. The results showed that no significant differences in fluency were identified between planning groups for either simple or complex tasks, nor between the high- and low-anxiety groups under each planning condition. These findings indicate that students' fluency was not significantly affected by either the anxiety level or the planning type.

5.3.5 Effects on adequacy

The findings for the effects of language anxiety, cognitive-task variability and planning type on students' adequacy are reported in this section. The subjective ratings on content, organization and overall text quality were adopted to quantify students' adequacy.

5.3.5.1 Effects on content

Table 5.36 shows the descriptive statistics of the simple and complex tasks across the six groups. The assumption of normal distribution was tenable, and the homogeneity of variances assumption was not violated for either the simple or complex task (Levene's *F* tests for content score in the simple task, F (5, 102) = 1.113, p = .358, and for content score in the complex task, F (5, 102) = .741, p = .595). A mixed ANOVA was, hence, conducted to assess the impact of language anxiety, cognitive-task variability and planning type on participants' score of content.

Table 5.36 Descriptive Statistics of Two Tasks across Six Groups for Content

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
Con-CG-H	18	18.39	1.84	17.97	1.66
Con-CG-L	18	19.42	1.91	18.61	1.59
Con-OLP-H	20	18.10	1.01	17.63	1.22
Con-OLP-L	16	18.78	1.48	18.31	1.29
Con-PTP-H	18	18.11	1.23	18.28	1.30
Con-PTP-L	18	18.89	1.76	19.92	1.70

Note. Con=content, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA (see Table 5.37) indicated the main effect for anxiety, F (1, 102) = 12.610, p = .001, η_p^2 = .110 and the interaction between task and group, F (2, 102) =

7.142, p = .001, η_p^2 = .123 were significant. Neither the main effect for task and for group, nor the interaction for task*anxiety, group*anxiety and task*group*anxiety were identified as significant. The results indicate that participants' content performance was markedly influenced by anxiety, if the cognitive-task variability and planning type were not considered. The content of participants in the simple and complex tasks differed across the PTP, OLP and CG conditions.

Table 5.37 Mixed ANOVA for Effects of Task, Anxiety and Planning on Content

	F	Sig.	η_p^2
task	1.282	.260	.012
group	1.852	.162	.035
anxiety	12.610	.001**	.110
task*group	7.142	.001**	.123
task*anxiety	.311	.578	.003
group*anxiety	.371	.691	.007
task*group*anxiety	1.669	.193	.032

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples t-tests identified a nearly significant difference in content between the simple (M = 18.89, SD = 1.76) and complex tasks (M = 19.92, SD = 1.70) in the PTP-L group t (17) = -2.101, p = .051, d=-.594, with a small to medium effect size. No significant differences in students' content performance were found for the other five groups between the simple and the complex tasks. These suggest that increases in cognitive-task variability slightly enhanced content for students with a low-level of anxiety and with pre-task planning before writing.

Independent-samples t-tests found that for the complex task, anxiety levels had a significant impact on PTP students' writing performance in terms of content: high anxiety (M = 18.28, SD = 1.30) versus low anxiety (M = 19.92, SD = 1.70), t (34) = -3.252, p = .003, d=-1.084 with a large effect size. Additionally, in the complex task, the content of PTP-L group (M = 19.92, SD = 1.70) was significantly higher than for the OLP-L group (M = 18.31, SD = 1.29), t (32) = -3.069, p = .004, d = -1.055 and the effect size was large. These findings suggest that when students completed the complex task, a low level of anxiety, compared to a high level of anxiety, enhanced PTP students' content. Pre-task planning, in comparison with online planning, also enhanced low-anxiety students' content.

5.3.5.2 Effects on organisation

Table 5.38 shows the descriptive statistics of the simple and complex tasks across the six groups. As the table illustrates, the low-anxiety students achieved a higher organisation score than their high-anxiety counterparts under each planning condition for both the simple and complex tasks.

Table 5.38 Descriptive Statistics of Two Tasks across Six Groups for Organisation

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
Org-CG-H	18	13.97	1.55	13.56	1.85
Org-CG-L	18	14.72	1.88	14.33	1.56
Org-OLP-H	20	13.83	.94	13.28	1.44
Org-OLP-L	16	14.16	1.29	13.59	1.07
Org-PTP-H	18	13.47	1.59	14.06	1.37
Org-PTP-L	18	14.78	1.52	15.33	1.80

Note. Org=organisation, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The assumption of normality for organisation was evaluated and found to be satisfied. Levene's *F* test for organisation in the simple task, F (5, 102) = 1.483, p = .202, and in the complex task, F (5, 102) = 1.446, p = .214, was not significant, indicating that the variance of the dependent variable in either the simple or the complex task was equal across the groups. A mixed ANOVA was, hence, conducted. As shown in Table 5.39, no interactions of task*anxiety, group*anxiety, and task*group*anxiety were identified for organisation. A significant interaction effect for task*group, F (2, 102) = 5.712, p = .004, η_p^2 = .101, however, was found that participants' organisation performance in the simple and complex tasks differed across the PTP, OLP and CG conditions. Only a main effect for anxiety was found, F (1, 102) = 9.925, p = .002, η_p^2 = .089, indicating that anxiety level substantially affected students' organisation, if cognitive-task variability and planning type were ignored.

Table 5.39 Mixed ANOVA for Effects of Task, Anxiety and Planning on Organisation

	F	Sig.	η_p^2
task	.775	.381	.008
group	2.597	.079	.048
anxiety	9.925	.002**	.089
task*group	5.712	.004**	.101
task*anxiety	.000	.989	.000
group*anxiety	1.229	.297	.024
task*group*anxiety	.003	.997	.000

Note. Task = cognitive-task variability, group = planning type.
**p<.05; ***p<.001

Paired-samples *t*-tests were performed, but no significant differences were found across the six groups between the simple and complex tasks, suggesting that cognitive-task variability had little effect on students' organisation across the six conditions.

A series of independent-samples *t*-tests was conducted, but there were no significant differences between high- and low-anxiety groups under each planning condition. For the complex task, organisation for the PTP-L group (M = 15.33, SD = 1.80) was better than for the OLP-L group (M = 13.59, SD = 1.07), t (28.1) = -3.472, p = .002, d = -1.159[21]. These findings indicate that pre-task planning, in comparison with online planning, enhanced low-anxiety students' organisation.

5.3.5.3 Effects on overall writing quality

Table 5.40 displays the descriptive statistics of the overall writing score, indicating that the low-anxiety groups outperformed the high-anxiety groups under each planning condition. To assess the impacts of language anxiety, cognitive-task variability and planning type on students' overall writing performance, the assumption of normality was first checked and found to be satisfied. Levene's *F* test for total writing score in the

[21] Levene's test for organisation in the complex task was significant, p=.030, values in *Equal variances not assumed* were reported.

simple task, F (5, 102) = 1.767, p = .126, and in the complex task, F (5, 102) = 2.174, p = .063, showed that the homogeneity of variances assumption was not violated for either the simple or complex task. Therefore, a mixed ANOVA was applied.

Table 5.40 Descriptive Statistics of Two Tasks across Six Groups for Overall Writing Quality

Variables	N	Simple task		Complex task	
		M	*SD*	*M*	*SD*
WQ-CG-H	18	64.08	5.82	61.83	6.19
WQ-CG-L	18	66.36	5.65	65.25	5.56
WQ-OLP-H	20	63.10	2.92	61.88	4.83
WQ-OLP-L	16	65.72	3.67	64.38	3.18
WQ-PTP-H	18	61.03	4.87	62.28	4.34
WQ-PTP-L	18	65.50	5.48	67.08	6.44

Note. WQ=writing quality, CG-H = high anxiety students under comparison group, CG-L = low anxiety students under comparison group, OLP-H = high anxiety students under online planning group, OLP-L = low anxiety students under online planning group, PTP-H = high anxiety students under pre-task planning group, PTP-L = low anxiety students under pre-task planning group.

The mixed ANOVA identified (see Table 5.41) significant main effects for anxiety, F (1, 102) = 14.777, p <.001, η_p^2 = .127, but not for task and group. The interaction between task and group, F (2, 102) = 5.057, p = .008, η_p^2 = .090, was significant, while interactions for task*anxiety, group*anxiety and task*group*anxiety were not significant. The results suggest that participants' writing quality was substantially influenced by anxiety, if cognitive-task variability and planning type were ignored. Writing quality in the simple and complex task versions significantly differed across the CG, PTP and OLP groups.

Table 5.41 Mixed ANOVA for Effects of Task, Anxiety and Planning on Overall Writing Quality

	F	Sig.	η_p^2
task	1.418	.236	.014
group	.172	.842	.003
anxiety	14.777	.000***	.127
task*group	5.057	.008**	.090
task*anxiety	.271	.604	.003
group*anxiety	.558	.574	.011
task*group*anxiety	.180	.836	.004

Note. Task = cognitive-task variability, group = planning type.
**$p<.05$; ***$p<.001$

A series of paired-samples *t*-tests was conducted to detect the differences between participants' total writing score on the simple task versus the complex task versions under the six conditions. The findings revealed that the only difference in overall writing quality to reach statistical significance was the CG-H group, in which participants performed significantly better in the simple task (M = 64.08, SD = 5.82) than in the complex task (M = 61.83, SD = 6.19), t (17) = 2.393, p = .029, d = .373. The effect size was small. These findings indicate that CG-H students' overall writing quality reduced significantly from the simple to complex tasks, as a function of increasing cognitive-task variability.

Independent-samples *t*-tests were conducted to compare the differences between the high- and low-anxiety groups under and between planning conditions. The results showed that students' overall writing scores in the PTP-L group (M = 65.50, SD = 5.48) almost significantly exceeded that in the PTP-H group (M = 61.03, SD = 4.87), when participants completed the simple task version, t (34) = -2.587, p = .014, d=-.862. The effect size was large. Similar differences were evident when pre-task planners completed the complex task: high anxiety (M = 62.28, SD = 4.34) versus low anxiety (M = 67.08, SD = 6.44), t (29.79) = -2.626, p = .014, d = -.875[22], with a large effect size. The findings reveal that a low level of anxiety contributed to students' overall writing

[22] Levene's test for overall writing quality in the complex task was significant, p=.036, and values in *Equal variances not assumed* were reported.

quality for both the simple and complex tasks, when students engaged in pre-task planning before writing.

5.3.5.4 Summary

In this section, the effects of language anxiety, cognitive-task variability and planning type on adequacy are summarised.

No significant interactions for task*group*anxiety, anxiety*task, or anxiety* groups were found in content, organisation and overall writing quality. Interactions for task*group and main effects for anxiety were found in the three measures of adequacy, indicating that participants' writing adequacy in the simple and complex task versions significantly differed across the CG, PTP and OLP groups, and that adequacy was influenced markedly by anxiety, if the cognitive-task variability and planning type were not considered.

Table 5.42 Summary of Significant Within-subjects Differences for Text Quality

Anxiety	Planning	Measures	Differences	Sig.	*d*
High anxiety	CG	WQ	Simple > Complex	.029	.373
Low anxiety	PTP	Content	Simple < Complex	.051(almost)	.594

The within-subjects differences across the three planning conditions and two anxiety groups are summarised in Table 5.42. For high-anxiety students, increasing cognitive-task variability had a negative effect on their overall writing performance under the CG condition. By contrast, for the low-anxiety group, students' content in writing benefited from pre-task planning before writing as well as raising the level of cognitive-task variability.

Table 5.43 Summary of Significant Between-subjects Differences for Text Quality

Task	Measures	Differences	Sig.	*d*
Simple	WQ	PTP-H < PTP-L	.014 (almost)	-.862
Complex	Content	PTP-H < PTP-L	.003	-1.084
		OLP-L < PTP-L	.004	-1.055
	Organisation	OLP-L < PTP-L	.002	-1.159
	WQ	PTP-H < PTP-L	.014 (almost)	-.875

The between-subjects differences under both the simple and complex tasks are summarised in Table 5.43. With the simple task, low-anxiety students' writing under the PTP condition was better overall than that of their high-anxiety counterparts. The same pattern was also found for the complex task: The PTP-L group outperformed the PTP-H group in overall writing performance, with the effect size larger than that in the simple task. The content of students in the PTP-L group was significantly better than that for the PTP-H and OLP-L groups with large effect sizes; similarly, with regard to organisation, the PTP-L group achieved higher scores in organisation than the OLP-L group in the complex task.

5.3.6 Summary

This section summarises the statistically significant results reported from 5.3.1 to 5.3.5, covering the findings for all indices of students' EFL writing performance.

Interactions for task*group*anxiety on writing performance were not identified, but main effect for anxiety was found in MLT, C/T, DC/C, MLC, CN/C, fluency, content, organisation and overall writing quality.

The within-subjects differences across the three planning conditions and two anxiety groups are presented in Table 5.44. For the high-anxiety group, increasing cognitive-task variability led to a significant reduction in students' C/T, DC/C and WT/√2W under the OLP condition, MLT, DC/C and CN/C under the PTP condition, and LD, WT/√2W and WQ under the CG condition. By contrast, increasing cognitive-task variability substantially enhanced low-anxiety EFL students' CP/C under the OLP condition, content and fluency under the PTP condition, and accuracy performance (fewer Err/T) under the CG condition. Of note, high-anxiety students' accuracy (i.e., Err/T, EP100 and EFC/C) and fluency, under the PTP group, improved as a function of raising the cognitive level of the task.

Table 5.44 Summary of Significant Within-subjects Differences for Students' EFL Writing

Anxiety	Planning	Measures	Differences	Sig.	*d*
High anxiety	OLP	C/T	Simple > Complex	.049	.559
		DC/C	Simple > Complex	.048	.556
		WT/√2W	Simple > Complex	.044	.506
	PTP	MLT	Simple > Complex	.001	.759
		DC/C	Simple > Complex	.027	.760
		CN/C	Simple > Complex	.032	.731
		Err/T	**Simple > Complex**	**.001**	**.944**
		EP100	**Simple > Complex**	**.029**	**-.498**
		EFC/C	**Simple < Complex**	**.048**	**-.503**
		Fluency	**Simple < Complex**	**.002**	**-.616**
	CG	LD	Simple > Complex	.010	.496
		WT/√2W	Simple > Complex	.000	1.353
		WQ	Simple > Complex	.029	.373
Low anxiety	OLP	CP/C	Simple < Complex	.047	-.703
	PTP	Content	Simple < Complex	.051(almost)	.594
		Fluency	Simple < Complex	.003	-.513
	CG	Err/T	Simple > Complex	.042	.537

In Table 5.45, statistically significant comparisons between the high- and low-anxiety groups across the OLP, PTP and CG conditions for both the simple and complex tasks are presented. Pre-task planners with low anxiety made better performance in WT/√2W and overall writing quality in the simple task than those with high anxiety, and PTP group's MLT, WT/√2W, content and overall writing quality for the complex task likewise benefited from a low level of anxiety. These patterns were, however, not identified for either the OLP or the CG conditions.

Table 5.45 Summary of Significant Differences between High- and Low-anxiety Groups for Students' EFL Writing

Task	Measures	Differences	Sig.	*d*
Simple	WT/√2W	PTP-H < PTP-L	.000	-1.674
	WQ	PTP-H < PTP-L	.014 (almost)	-.862
Complex	MLT	PTP-H < PTP-L	.003	-1.077
	WT/√2W	PTP-H < PTP-L	.001	-1.322
	Content	PTP-H < PTP-L	.003	-1.084
	WQ	PTP-H < PTP-L	.014 (almost)	-.875

Finally, the statistically significant differences among the CG, OLP and PTP groups under each anxiety level for both the simple and complex tasks are summarised. As Table 5.46 displays, in the simple task, students in the PTP-H group did not perform as well as those in the CG-H group in terms of WT/√2W and accuracy, and those in the OLP-H group in terms of WT/√2W; no significant differences among the CG, OLP and PTP groups were observed for the low-anxiety students in the simple task. Interestingly, the opposite pattern was found in the complex task: No significant differences were identified across the CG, OLP and PTP groups when students were highly anxious. The low-anxiety students, in the OLP (in MLC and CP/C), and PTP groups (in MLC), had greater scores for syntactic complexity than the CG when completing the complex task. Students in the PTP-L group significantly outperformed those in the OLP-L group in the complex task in content and organisation performance, while the CG-L group wrote more accurately than the OLP-L group in the complex task.

Table 5.46 Summary of Significant Differences among OLP, PTP and CG under High- and Low-anxiety Groups for the Simple and Complex Tasks

Task	Measures	Differences	Sig.	*d*
Simple	WT/√2W	PTP-H < CG-H	.000	-1.894
		PTP-H < OLP-H	.000	-1.579
	Err/T	PTP-H > CG-H	.007	.953
	EP100	PTP-H > CG-H	.009	.931
	EFC/C	PTP-H < CG-H	.002	1.108
Complex	MLC	CG-L < OLP-L	.005	-1.045
		CG-L < PTP-L	.013	-.878
	CP/C	CG-L < OLP-L	.001	-1.275
	Err/T	OLP-L > CG-L	.009	.955
	EP100	OLP-L > CG-L	.012	.915
	Content	OLP-L < PTP-L	.004	-1.055
	Organisation	OLP-L < PTP-L	.002	-1.159

5.4 Perceptions of Anxiety Related to Different Conditions

The results for RQ3c and RQ3d are reported in this section. The purpose of the qualitative phase was to explain and complement the quantitative data and provide deep insights into what participants with different anxiety levels do and feel when they engage in different conditions. Following the quantitative phase, 15 students were purposively recruited on a voluntary basis for individual interviews. There were 3 females in the OLP-H group, 2 females in OLP-L, 3 females in PTP-H, 2 males in PTP-L, 1 male and 1 female in CG-H and 3 females in CG-L. The data obtained from the interviews were analysed qualitatively with both bottom-up and top-down approaches employed. To present the findings, participants' direct quotes were translated verbatim from Chinese to English so that their ideas would not be altered, with representative quotes selected by the researcher to exemplify the issues that arouse and the points made. The interview questions are in Appendix E.

5.4.1 Perceptions and sources of writing anxiety

In the individual retrospective interviews, the participants were asked to evaluate their self-perceptions of writing anxiety when they wrote an English composition (Interview Q1: Do you feel anxious or not when you write English compositions? Why?). The 8 participants (100%) from the high-anxiety group confirmed that they experienced a high level of anxiety when they wrote an English composition. Some participants acknowledged their highly anxious feelings for examinations, while others stated that their anxiety was aroused when they were asked to write compositions in English.

> "I feel nervous in practice and even more anxious in timed writing."—*CG-H-02*
>
> "Whenever I write English compositions (no matter it is an exercise or an examination), I feel anxious."—*OLP-H-04*
>
> "I am a little nervous when I write English compositions for homework or exercise by myself. For timed examinations, the anxiety will double."—*PTP-H-06*.

Further analyses of their interview responses showed that self-perceived low writing proficiency, time constraint, task difficulty, conceiving anxiety, and differences between Chinese and English were the factors high-anxiety students identified as contributing to their anxiety in writing in a second language (see Table 5.47). No single source of anxiety was mentioned by the high-anxiety participants.

Among the reasons listed in Table 5.47, high-anxiety students most frequently mentioned time and attributed their substantial anxiety to time constraints. Six of the 8 high-anxiety participants reported that they suffered from a high level of anxiety due to time limits for writing. For 4 students, the situation worsened when they did not finish writing in the last minutes, as seen in the following excerpts.

> "When time is up and I do not finish writing, I will feel panic and rush to end it up (reach the word limit) without considering the quality."—*CG-H-01*
>
> "If I did not finish in the last 5 minutes, my hands would shiver and I would feel very panic. At that time, I cannot find a word to express my thoughts."—*OLP-H-03*

Table 5.47 Sources of Second Language Writing Anxiety for High-anxiety Group

Sources of writing anxiety		Number	Frequency (%)
Time constraint:		**12**	**54.55%**
	Time limit	6	
	The last minutes	4	
	Slow writing speed	2	
Self-perceived low writing proficiency		**4**	**18.18%**
Conceiving anxiety		**3**	**13.64%**
Task difficulty		**2**	**9.09%**
Difference between Chinese and English		**1**	**4.55%**
Total		**22**	**100%**

Another important reason mentioned by participants in interviews was lack of confidence in their writing proficiency. Some participants said they felt anxious when writing in English, because they cannot accurately express their thoughts appropriately in English; some interviewees expressed concerns that they can use only simple words and sentence patterns instead of native English and complex expressions in English.

Participants from the low-anxiety group differed from their high-anxiety counterparts. Four out of the 7 participants implied that they did not feel anxious when they were writing English compositions, and the other 3 stated that they felt slightly nervous in some certain situations. Specifically, #13 (OLP-L) and #15 (PTP-L) reported that they simply felt a little bit nervous in some formal and important examinations; #10 (CG-L) said: “I usually don’t feel anxious except for the philosophic topic, which is very abstract and difficult to understand”.

In summary, consistent with the results from the SLWAI, the high-anxiety interviewees expressed experiencing a higher level of writing anxiety than their low-anxiety counterparts.

5.4.2 Anxiety related to cognitive-task variability

In the retrospective interviews, participants were asked to recall their self-perceptions on the task difficulty and anxiety feelings on the two tasks.

Seven out of the 8 high-anxiety interviewees indicated that the complex task[23] was more difficult than the simple one because of the number of elements involved, more mental effort required for both organisation and content, and a heavier language load. Only one participant, #02 (CG-H) was unable to choose which task was more difficult, as she found both tasks challenging.

Resonating with the concerns of the high-anxiety participants, all the low-anxiety interviewees characterized Task 2 as more complex than Task 1. Interestingly, #13 (OLP-L) vividly gesticulated a tiny distance (i.e., showing a 1 cm distance with her fingers), explaining that the gap of the task difficulty between the two tasks was very small, although Task 2 was a little bit more difficult than Task 1. The reasons provided by the low-anxiety participants were similar to those given by their high-anxiety counterparts. One of the low-anxiety participants, #14 (PTP-L) said that he found Task 2 more interesting and appealing.

The interview responses of both high- and low-anxiety participants to the task difficulty question confirmed the manipulation of the cognitive-task variability of the two tasks resulted in a differentiation of the cognitive load, as determined in the pilot study (see Section 3.4.2) and in Phase I (see Section 4.2).

In response to the question: Which task made you more anxious, 5 out of the 8 high-anxiety participants reported that the complex task made them more anxious, and that task difficulty mainly contributed to their higher levels of anxiety, as the following excerpts illustrate.

> "On the one hand, the two pairs of roommates I chose in Task 2 have some similarities, which easily makes repetitions, and on the other hand, the differences between the two pairs should be written in detail. How to organize the essay and make a balance make me anxious."—*CG-H-01*

> "I cannot choose the perfectly matched roommates in Task 2, which takes me more time to conceive the structure of the composition. These reasons arouse my anxiety."—*OLP-H-03*

[23] In order to avoid the potential effects of the task name on students' responses, in the interview Task 1 refers to the simple task and Task 2 refers to the complex task.

> "I wrote Task 1 for the first time. Compared to Task 1, I found Task 2 is more difficult. I did not do well in Task 1, so I felt more nervous when I wrote Task 2."—*OLP-H-05*

> "Task 2 is more difficult. It requires you to read, think and output in a more quick way. Time was not enough and I ended Task 2 in a hurry."—*PTP-H-08*

The remaining 3 high-anxiety interviewees unexpectedly expressed that they felt more nervous when they completed Task 1. They gave a somewhat unitary reason for their feelings that they completed Task 1 for the first writing time. Recall that the two tasks were counterbalanced: Half of the students completed Task 1 first and half Task 2 first. For these students, the unfamiliar task type made them nervous. For example, #04 from OLP-H summarized her feelings:

> "I wrote Task 1 for the first time and Task 2 for the second time. It is the first time for me to write such kind of task. I mean it is not like the argumentative compositions I wrote before which often give you two viewpoints and ask you to choose the one you agree and give your reasons. I am quite unfamiliar with writing this type of task which gave me lots of options and information in a table—I did not know how to start and where to start. You know Task 1 and Task 2 share a similar structure. When I wrote Task 2 at the second time, I knew the task (I mean the structure of this kind of task) and I did not feel as nervous as I wrote Task 1 at the first time, though Task 2 was more difficult and I did not remember what I previously wrote in Task 1. It's a sense of safety."

For the 7 low-anxiety students, the responses were interesting and varied. Three out of the 7 (43%) participants reported that they felt more anxious completing the complex task due, primarily, to task difficulty and mental effort. Two participants stated that the simple task contributed more to anxious feelings, because of the sequencing and unfamiliarity of the task. Another two, #11 and #13, indicated that they did not feel nervous about either simple or complex task, as the following excerpt illustrates:

> "I think I can structure the essays in a very quick way, when I wrote English compositions. I wrote according to the structure I listed, so I did not feel anxious in either Task 1 or Task 2."—*CG-L-11*

5.4.3 Anxiety related to planning type

During the interviews, students in each of the groups, pre-task planning, online planning and comparison groups, were asked on what they focused during planning and writing.

Based on their responses, and on previous research (Manchón & De Larios, 2007), participants' focus of planning and writing was classified into three broad categories: Content, organisation and language. The categories and operational definitions are presented in Table 5.48, with examples of interviewees' responses by category.

Table 5.48 Categories for Students' Focus during Planning and Writing

Category	Definition	Samples
Content	Ideas generated by the writer that will be subsequently formulated in a linear fashion or discarded	*"During the pre-task planning, I will first make a decision on the two pairs of roommates."—PTP-H-07* *"I will list the similarities and differences between the two candidates."—PTP-L-15* *"I tried to figure out the reasons for choosing these two candidates as roommates."—OLP-H-05*
Organisation	Decisions about how the ideas in the text are to be framed or sequenced in connection with the text type in hand and its rhetorical structure: introductions, conclusions, examples, paragraphs.	*"I will make an outline to figure out logic and structure…"—PTP-H-08* *"My main focus during pre-task planning is structure and content."—PTP-L-14*
Language	The linguistic encoding of ideas in terms of lexical complexity (vocabulary), syntactic complexity (sentence structures), and accuracy (grammar).	*"When I write, I will continue thinking. The main focus is language expression, and sometimes the connection between two paragraphs or ideas."—PTP-H-07* *"I mainly focused on language, such as word choices, sentence patterns, and so forth."—OLP-H-04*

Note. The definition of Content, Organisation and Language are adopted from (Manchón & De Larios' study (2007).

The categories cover the answers in which the words "content, organisation or language" were explicitly mentioned (e.g., my main focus is content) and those that implicitly indicated a certain focus (e.g., "I made the decision for the roommates first"). It is acknowledged that a single statement may contain more than one focus and that the three categories are sometimes interwoven in students' statements. In such cases, the statement was coded according to the category represented in the statement. A single statement was, therefore, coded as more than one category. For example: "I will make an outline to figure out the logic, structure and content" (PTP-H-08). In this case, the focus was coded as "organisation" and "content", as the words "logic and structure" implicitly indicated the category "organisation" and the word "content" was explicitly mentioned.

5.4.3.1 Pre-task planning group

Participants in the PTP group were instructed to do the pre-task planning in 15 minutes and then write the essay in 25 minutes. Two of the 5 interviewees reported that they stopped to think during writing for a few seconds; one referred to language expression and the other mentioned recalling the outline listed in the pre-task planning. All 5 commented that they had no time left to edit the essays. These indicate that the online planning they engaged in was limited during writing, which is consistent with the research design of this study.

5.4.3.1.1 Focus and strategies

As shown in Table 5.49, the focus during pre-task planning and during writing as mentioned by interviewees in the PTP group was different. Students mainly focused on content (55.56%) and organisation (33.33%) when they did pre-task planning, but during the writing process, the focus on language (66.67%) was mentioned most frequently.

Table 5.49 Planning and Writing Focus Reported by Pre-task Planners

Focus	**Pre-task planning**		**Writing**	
	Raw	**Frequency**	**Raw**	**Frequency**
Content	**5**	**55.56%**	**2**	**22.22%**
Organisation	**3**	**33.33%**	**1**	**11.11%**
Language:	**1**	**11.11%**	**6**	**66.67%**
Vocabulary	1		2	
Sentence	0		0	
Grammar	0		1	
Language (general)	0		3	
Total	9	100%	9	100%

No big differences in the foci during planning and writing were observed between the high-anxiety participants and their low-anxiety counterparts. The specific strategies they adopted, however, according to their interview responses, vary greatly.

The high-anxiety participants reported trying to reach a "hundred percent preparation" during their pre-task planning. Two out of the 3 reported that they wrote drafts of the two tasks in the 15-minute planning time:

> "I wrote rough drafts in both English and Chinese."—*PTP-H-06*
>
> "I made the decision first. Then, I wrote the first draft in English before writing. We did not learn how to do pre-task planning in the writing courses, and I usually make a draft when writing in English."—*PTP-H-07*

In terms of writing, pre-task planners with high anxiety appeared to transfer what they had pre-planned with little development. For example, they used simple words and sentences that first came to their mind to express their pre-planned ideas, or translated their thoughts from Chinese to English. One of the high-anxiety participants, #06, emphasized that grammar (i.e., accuracy) was her primary focus during writing.

In contrast, the low-anxiety participants did the pre-task planning in a more global way. Both the interviewees reflected that they did organizational planning by making an outline and rehearsed by listing a few details they were supposed to use. When engaging in the writing process, low-anxious participants used some writing strategies, such as elaboration, although they shared a similar language focus during writing, with their highly anxious counterparts. For example, #15 stated that he added more details into his pre-planned arguments to support statements when writing and #14, a male participant, stated that he introduced additional connective words to enhance the coherence of the essays.

5.4.3.1.2 Perceptions of pre-task planning

The interview data reflect a uniform pattern of the awareness that pre-task planning had definitely improved their writing fluency. For other functions of pre-task planning, however, the high- and low-anxiety participants did not concur.

The 3 highly anxious participants said they benefited from pre-task planning in varying respects. For example, pre-task planning helped them organize their thoughts (100%), reduce their anxious feelings (100%), write information relevant to the topics (33.33%), and prepare well for vocabulary (33.33%).

A high-anxiety female participant, #08, compared her experience of writing these two tasks with her previous writing experience. A high level of anxiety usually pushed her to do the pre-task planning in a rushed manner, which offered little help to her writing in terms of writing quality and affective state. Compared to her previous fast-paced planning activities, it seems that she prepared well for writing the two tasks both physically and mentally during the 15-minute pre-task planning. She said:

> "For every examination or writing test, I will leave enough time for writing, but I still feel very nervous. After reading the writing prompts, it usually takes me 2-3 minutes to do the pre-task planning, like thinking about the structure and content. I used to control the pre-task planning time within 5 minutes, because I am very worried about I could not finish writing within the time limit. You know, sometimes, 2-3 minutes is not enough for me to figure out everything clearly, but I have to start writing, no matter whether I figure everything out or not. I have to admit that in most cases, I cannot figure it out in a very clear and logical way. For these two tasks, I was asked to plan in 15 minutes and write in 25 minutes, which gives me enough time to do the pre-task planning. For these two tasks, I thought about the structure and content thoroughly and clearly and made a much clearer outline. I think I write more fluently than before. Pre-task planning, to some extent, increases my confidence and reduces my anxiety."

Apart from the advantages of the pre-task planning, #08 also admitted that the transfer of the plan into on-line performance, in particular, the language part, was hindered by the limited writing time: “I have to use simple language to express my ideas”; and “What I had planned was not fully transferred into the final writing text. It’s hard to remember everything”.

Not all participants appeared to benefit from pre-task planning equally. Participant #07 believed that PTP did offer some help, but not enough. She found the anxious feelings were raised when she had to transfer what she had planned within limited time during the writing process, although planning prior to the task, to some extent, eased her tension. She stated that she would prefer her own writing habits in her later writing, using both PTP and OLP instead of only PTP.

The degree of benefits of pre-task planning for the two tasks, even for the same participants, was not equal. As Subject #06 reflected “at the first time, I was very nervous, because it is the first time for me to write such kind of task. I was very nervous and I did not plan as well as what I planned for the second time. And I did not make very good use of what I had planned, compared to the second time”. From her comments, it seems that anxiety could affect the quality of pre-task planning and, then, the writing to some extent.

The picture is different for the low-anxiety participants. Compared to their high-anxiety counterparts, participants with a low level of anxiety reported less benefit from pre-task planning. One of the low-anxiety participants unexpectedly acknowledged that 15-minute pre-task planning, to some extent, increased anxious feelings a bit because “he was kind of worried that the writing time (25 minutes) was not enough”.

5.4.3.2 Online planning group

Participants in the online planning group were asked to start writing immediately after reading the writing task prompts. In this case, they were supposed to do online planning while they were writing. Four of the 5 online planners interviewed, reported they first spent a little time reading the instructions, quickly understanding the task requirements, and then started writing. Only one interviewee acknowledged that she made a general plan in her mind (maybe 5 minutes) after understanding the prompts and before beginning to write. None reported they had time to edit the essays. From the responses, it is apparent that the pre-task planning was limited and online planning occurred during writing; this was the intention of the research design of this study.

5.4.3.2.1 Focus during online planning

Information on the students' focus while writing, as well as planning, was gathered through the retrospective interviews. As shown in Table 5.50, overall, the main focus for the online planners was language (66.67%), followed by content (26.67%) and a small proportion of time spent on organisation (6.67%).

Table 5.50 Online Writing Focus Reported by Online Planners

Focus	**High anxiety**		**Low anxiety**		**Total**	
	Raw	**Frequency**	**Raw**	**Frequency**	**Raw**	**Frequency**
Content	**3**	**37.5%**	**1**	**20%**	**4**	**30.77%**
Organisation	**1**	**12.5%**	**0**	**0**	**1**	**7.69%**
Language:	**4**	**50%**	**4**	**80%**	**8**	**61.54%**
Lexical complexity	1		2		3	
Syntactic complexity	1		2		3	
Accuracy	0		0		0	
Language (general)	3		0		3	
Total	**8**	**100%**	**5**	**100%**	**13**	**100%**

For the high-anxiety group, all participants paid attention to two aspects, language and content; one interviewee also focused on organisation. Neither the explicit words, "accuracy" and "grammar", nor the implicit expressions of accuracy were detected in the interview data, although most attention was paid to the language category. Most described their online planning of language generally without any explanation, such as "language expression is my main focus". Only one participant stated that she carefully considered word choices and sentence structures during her writing.

Interviewees from the low-anxiety group seemed to lack concern about organisation and language accuracy during their online planning process. Similar to their high-anxiety counterparts, the low-anxiety online planners mainly focused on the language. The 2 low-anxiety interviewees, unlike the high-anxiety participants, reported they had a specific focus on language complexity, acknowledging that they tried to use more complex words and sentence patterns, as the following excerpts illustrate:

> "I tried to search for some complex sentence patterns and sophisticated vocabulary, such as the words with more than three syllables, in my mind."—#12

> "During writing, I mainly focused on language expression. I tried to use some complex sentence constructions and native expressions."—#13

5.4.3.2.2 Perceptions of online planning

Responses on the 5 interviewees' attitudes towards online planning fell into three patterns: Some online planners (2 high- and 1 low-anxiety students) expressed preference for their own writing habits (i.e., a short PTP + OLP), rather than simply planning online in their later writing activities; one high-anxiety participant indicated that while she may use online planning when the writing time is insufficient, she would usually use a PTP+OLP pattern; a low-anxiety participant stated that the availability of only online planning did not make a difference to how she felt about the task and her performance success.

The high-anxiety participants provided more seemingly negative evaluations of the online planning during the interview. All highly anxious participants complained that with only online planning, their writing anxiety increased. Two participants (#3 and #5) reported that during writing, they got stuck more frequently; wrote the content less logically and less coherently; and could not think deeply when starting writing immediately, although they had enough time to do online planning during writing, as the following excerpts illustrate:

> "I, again and again, went back to the sentence I have written before"—#05

> "Sometimes, I even did not know where I was when I wrote"— #05

> "So bad! What I wrote was not logical."—#03

Declining writing fluency, reduced organisation and poor content performance consequently led to a higher level of anxiety. For #04, a female high-anxiety participant, her antipathy to starting writing immediately was very strong, adversely influencing her feelings and intensifying her anxiety:

> "I was required to start writing the two tasks immediately without sufficient planning, which makes me more anxious. I tried to push myself to start immediately, but I was too nervous to start writing without a clear picture in my mind. I was kind of frozen in 1-2 seconds. I cannot do that, so I quickly made an outline in my mind in about 5 minutes."

Resonating with the concerns of the high-anxiety participants, one low-anxiety interviewee, #12, described OLP as making her a little bit uneasy. Online planning (start writing immediately) led her to think that what she had written before was not coherent with what she was writing now, with statements becoming overlapping or disconnected. The reasons she gave were consistent, to some extent, with those provided by the high-anxiety interviewees.

Interestingly, another low-anxiety participant realised that she experienced no big difference in the writing performance and affective feelings between the two writing tasks and her previous writing activities.

5.4.3.3 Comparison group

Students in the comparison group were instructed to complete the two writing tasks in their own writing style. Two participants said they started writing immediately and 3 adopted a PTP+OLP pattern when writing (the time allocated to PTP and OLP varied). Three of the 5 acknowledged that 2-5 minutes was available to edit after they finished writing, and the remaining 2 students finished writing on time. It appears that the comparison group differs from the PTP and OLP groups in the amount of time allocated to planning and writing.

5.4.3.3.1 Focus during writing

Overall, the most frequent focus areas (see Table 5.51), reported by the 5 CG participants, were language (44.44%) and organisation (33.33%). Students from both high- and low-anxiety groups, unlike the PTP and OLP groups, frequently mentioned a specific language focus, accuracy rather than complexity.

Regarding the high-anxiety group, #02 tried to reach "a 100% correctness" even at the price of complexity, as the following excerpt illustrates:

> "As for the focus, accuracy comes first. I wrote simple sentences instead of complex ones to make sure the accuracy".

Another high-anxiety interviewee, unexpectedly, thought her priority was to finish the tasks on time and she had no specific focus during her writing. When writing English essays, she was always in a hurry, quickly writing what came into her mind, because she was afraid that she may not have enough time to complete the tasks.

With the low-anxiety group, 1 of the 3 participants focused on accuracy, saying that "What I mainly focused on is how to express my ideas accurately". Apart from the accurate language expression, 2 other low-anxiety students paid most attention to organisation.

Table 5.51 Writing Focus Reported by Comparison Group

Focus	Raw	Frequency
Content	**1**	**11.11%**
Organisation	**3**	**33.33%**
Language:	**4**	**44.44%**
Lexical complexity	1	
Syntactic complexity	0	
Accuracy	3	
Task completion*	**1**	**11.11%**
Total	**9**	**100%**

Note. The focus with* is not in the three broad categories—content, organisation and language.

5.4.3.3.2 Writing preference in comparison group

In the interview data, the 2 high-anxiety participants both reported that they started writing after understanding the writing prompts. They said that they were too anxious about the time constraint to allocate extra time into pre-task planning and lacked confidence in completing the tasks within the time limit. And as a matter of fact, the CG condition, however, did not obviously help either the writing fluency or the affective feelings, as seen in the following excerpts.

> "I was stuck sometimes when I wrote. As I wrote, I found this statement was not suitable here. Then, I doubted whether I had written down a wrong statement, which made me

> more nervous. Even worse, the increasing anxiety made my mind go blank and negatively influenced my output, but I have to go ahead, no matter whether it was good or not. I mean I have to finish the writing first."—*CG-H-01*

> "As I wrote, I sometimes stopped to think, because, at that point, I don't know how to continue my essay writing. This aroused my anxiety. I have to push myself to calm down and continue writing." —*CG-H-02*

The two high-anxiety participants preferred this writing style simply because their priority was to complete the tasks within the time limit, instead of achieving successful writing.

By contrast, the interview data shows that the low-anxiety participants shared a similar writing pattern, PTP+OLP, although the time allocated to PTP and OLP varied. They preferred to do short pre-task planning, such as "making an outline", "listing the statement and supporting points" or "listing the three reasons for choosing the roommates". With the help of pre-task planning, students did not need to consider the statement when writing. Instead, they could pay more attention to the language "to express their ideas correctly", but also add more detailed content into the body of their writing.

Further analyses of the low-anxiety participants' interview data indicate that they all evaluated their writing pattern positively. For example, #09 stated that doing pre-task planning and adding details while writing helped her write more fluently. #10 attributed her low level of anxiety to a similar writing pattern.

> "I used to feel a little nervous at the beginning of writing. When I do pre-task planning, I focus on how to make an outline, which distracts my attention from being unease to pre-task planning".

Likewise, #11 supported her writing preference. "Even the time is not enough, for example, only 30 minutes left, I will do 5-to-6-minute pre-task planning when I write. During writing, I can add more detailed supporting points into the essay", which she thought helped her write more logically and coherently.

In comparing the two anxiety groups of participants, the level of anxiety, to some extent, appeared to influence students' preference for writing styles in their writing activities, and their writing patterns, in turn, affected their feelings and writing performance.

When asked how they dealt with the time remaining when they finished writing, 3 participants (1 high-anxiety and 2 low-anxiety students), who finished in advance, acknowledged that they used the time for editing grammar errors.

5.4.4 Summary

The above sections have presented the findings that were obtained from the interview, which show that the high-anxiety interviewees generally experienced an even higher level of writing anxiety than their low-anxiety counterparts. High-anxiety participants commented that their anxiety was provoked mainly by the time constraint and self-perceived low writing proficiency.

The interviews revealed that the relationship between anxiety, cognitive-task variability, and planning type was non-linear.

Most participants (87.5% high-anxiety and 100% low-anxiety students) characterized the complex task as more difficult and requiring the consumption of more mental effort than the simple task. One highly anxious student, who found both tasks difficult, was the only exception. Participants' anxious feelings towards the two tasks, however, did not reach consensus like their perceptions of task difficulty. More than half the participants (53.3%: 5 high- and 3 low-anxiety students) reported that they were more nervous when completing the complex task due to the higher level of task difficulty. Thirty-three percent of interviewees acknowledged that they felt more anxious when writing Task 1, and that the task sequencing and unfamiliar topic contributed to their nervous feelings. Two low-anxiety participants did not experience different levels of anxiety when they completed the two tasks.

The students reported focusing on different aspects under the three planning conditions. The pre-task planners said that during planning they mainly focused on content and organisation, and the main concern was for language use during the writing process. Highly anxious participants preferred a more detailed local planning style, such as drafting, while low-anxiety students planned in a more global way, such as organisational planning. High-anxiety students in the PTP group reported they benefited more from pre-task planning, in terms of reducing their anxiety, than their low anxious counterparts, but some high-anxiety interviewees noted that they felt pressured when transferring what they had planned into writing text due to the limited writing time.

With online planning, more interviewees focused on language than content and organisation. Low-anxiety online planners paid more attention to the complexity of language expression, unlike their high-anxiety counterparts, who described their online planning for language globally without detailed explanation. Most online planners commented negatively on online planning, such as “increasing anxiety” or “reducing writing fluency”; one low-anxiety participant, however, reported no major differences compared to her previous writing activities.

The CG participants focused on both language (in particular accuracy), and organisation. One of the 2 high-anxiety participants was primarily concerned with accuracy at the price of complexity, while the other prioritised task completion. The low-anxiety participants attached importance to organisation as well as language accuracy. In their approach to writing, the high-anxiety interviewees acknowledged that they started writing immediately and did not allocate extra time for pre-task planning because they had no confidence in finishing the tasks in the time allotted. In contrast, all the low-anxiety participants used a PTP+OLP writing pattern, and positively evaluated their writing style.

5.5 Chapter Summary

The results obtained from the quantitative data (Stage II of Phase II) and the qualitative data to address RQ3 in this chapter are found consistent. As reflected in the interviews, participants from the high-anxiety group seem to be more prone to L2 writing anxiety than their counterparts from the low-anxiety group. Interview responses of participants from high- and low-anxiety groups support the results identified through the *SLWAI*.

The quantitative findings mainly showed that anxiety did not statistically interact with cognitive-task variability and planning type to influence students’ writing performance, but interfered with the effects of cognitive task complexity on writing performance. To elaborate, the effects of cognitive-task variability on performance are influenced differently by anxiety with increases in cognitive task complexity resulting in a decrease in high-anxiety participants' writing performance, but an increase in low-anxiety participants' performance. Furthermore, the effects of planning type on performance differ between the high- and low-anxiety groups. For instance, in the low-anxiety group, the planning groups (i.e., the PTP and OLP groups) had higher achievement in syntactic complexity than the CG in the complex task, while the high-anxiety students in the PTP group had lower achievement in lexical complexity, compared to the CG and OLP groups. Low-anxiety students were at an advantage in writing performance over their high-anxiety counterparts under the PTP condition, which did not occur under the OLP and CG conditions.

The qualitative results confirmed that anxiety, cognitive-task variability, and planning type interacted to influence students' L2 writing performance. While different levels of cognitive-task variability and varying planning types, to some extent, arouse or reduce participants' anxious feelings, students' anxiety levels, more or less, influence their perceptions of cognitive-task variability and different planning conditions.

In the next chapter, the research findings presented in Chapter Four and Chapter Five are further interpreted and discussed.

Chapter 6 Discussion

6.1 Overview

This chapter discusses the results of the study in reference to relevant empirical studies as well as theoretical perspectives and frameworks. The discussion is divided into three main sections based on the three overarching research questions, focusing in turn on: (a) Validating the manipulation of cognitive-task variability; (b) interpreting the effects of cognitive task complexity along each and both the two dimensions on EFL writing performance; and (c) interpreting how writing anxiety interferes with cognitive task complexity in the field of EFL writing.

6.2 Validation of Cognitive-task Variability

This section discusses the research question: *What are the relationships between presumed cognitive-task variability and cognitive load?* The results for self-ratings and dual-task experiments are interpreted and discussed respectively.

6.2.1 Self-ratings

The results found that the designed-to-be complex task was perceived as significantly more difficult and requiring significantly more mental effort than the designed-to-be simple task version. These results lend support to Robinson's (2001b, 2005, 2011a) hypothesis that enlarging the number of elements increases the cognitive task complexity along the resource-directing dimension, leading to greater cognitive load. The findings of self-ratings on task difficulty and mental effort are in line also with the previous research (Révész et al., 2016; Sasayama, 2016). For example, in Révész et al.'s (2016) study, participants rated the designed-to-be complex task more difficult, consuming more mental effort when completing the map and decision-making tasks than the simple task. Sasayama (2016) found that among the four tasks, the designed-to-be simplest task was rated with the lowest level of cognitive-task variability and the most complex task with the highest level of cognitive complexity, measured by both self-ratings of cognitive load.

The findings in the present study, however, are inconsistent with aspects of Révész et al.'s (2016) study, in which no significant differences were reported between the participants' ratings of task difficulty on the narrative tasks. The difference in outcomes may be due to: different task modalities, different task genres or varying levels of cognitive load between the simple and complex tasks. For example, writing tasks were used in the present study while oral tasks were adopted in Révész et al.'s (2016) research, and while speaking is a linear process, writing is a recursive process (Kellogg, 1996). The difference in speaking and writing modalities might lead to participants' different perceptions of task difficulty and mental effort, since people have different schemata for different tasks, demands, or experiences (Flower & Hayes, 1984; Reed, Burton, & Kelly, 1985). Additionally, learners' schemata for completing tasks differ with genres, such as narrative, descriptive and persuasive tasks (Flower & Hayes, 1984; Reed, Burton, & Kelly, 1985). The argumentative genre used in the present study, compared to the narrative, appears to require increased cognitive demands (Andrews, 1995; Freedman & Pringle, 1984), more critical thinking and articulation of ideas, and greater challenge in the use of linguistic elements (Hirose, 2003). Learners may perceive the argumentative tasks as more difficult and mental effort consuming than the narrative tasks. Moreover, the difference in cognitive load between the simple and complex task versions might have been more detectable to participants, in terms of perceived task difficulty and mental effort, in the present study than that in Révész et al.'s (2016) research. The complex task in Révész et al.'s (2016) research, compared to the simple task, entailed an unpredictable set of unfamiliar events which required greater reasoning demands. In contrast, the complex task in the present study was characterized by more argument elements and reasoning demands, with two additional candidates and two additional properties of each candidate from which to choose, than the simple version. As Rahimi (2019) explained, internal cognitive comparisons will be increasingly imposed, when elements in the task are closely related and affect each other. In light of this, the complex task in the present study might be percieved as even more difficult than the simple one. These task manipulations, combined with the argumentative genre, may lead to a larger gap of cognitive load between the simple and complex tasks in the present study than that in Révész et al.'s (2016) research; as Sasayama (2016) speculated "cognitive complexity of tasks may be experienced differently only when the tasks differ from each other considerably in terms of number of elements" (p. 243).

Comparing the effect sizes of the two measures, although the results of the TD and ME measures revealed similar trends across the two writing tasks, the effect size of TD (d=.521) was larger than ME (d=.418). This finding suggests that the two variables, task difficulty and mental effort, are related but not identical. As previously discussed, task difficulty is based on learners' perceptions of the cognitive demands of the tasks, determined by the abilities and affective responses that learners bring to the task (Robinson, 2005, 2007a, 2011a), while mental effort is based on the actual effort engaged in the tasks. The larger effect size of TD than ME might suggest that whereas participants perceived the complex task as more difficult than the simple task, they exerted only a little more effort when completing the complex than the simple task versions. The findings of strong, but not perfect, correlations between perceived task difficulty and mental effort further confirm that the two variables are related but not isomorphic (Brünken et al., 2010; Révész et al., 2016).

6.2.2 Dual-task experiments

RQ1b addressed the issue of whether different intended levels of cognitive-task variability result in varying levels of cognitive load during task performance as measured by dual-task method. Two dual-task experiments with different versions of the secondary task were conducted in the present study. With dual-task Experiment 1, the results showed that participants had significantly higher accuracy rates on the secondary task when carrying out the simple primary writing task versus the complex one. This finding aligns with the previous studies (Révész et al., 2016, 2014), and partially echoes the principle underlying the dual-task paradigm that when the level of cognitive load imposed by the primary task is increased, less accurate and more reaction time are expected on the secondary task. The results for secondary-task accuracy, therefore, suggest that the more complex version of the primary writing task in the present study imposed greater cognitive demands.

No significant differences were identified for secondary-task reaction time, however, when participants were under the simple primary task condition, as compared to under the complex one. This finding is not surprising, since the same trends were observed in previous research (Révész et al., 2016, 2014; Sasayama, 2016). This is probably because the design of the secondary task in the present study is relatively similar to those used in previous studies (Révész et al., 2016, 2014; Sasayama, 2016), although the modality of the two secondary tasks is different (i.e., one is auditory task and others are visual task). Specifically, the secondary task in the present study required participants to detect the change of two sounds. The other three studies (Révész et al., 2016, 2014; Sasayama,

2016) asked students to notice the change of two colours and react to one of the colours. An explanation for the non-salient reaction time, as noted by Révész et al (2016), may be that the secondary-task reaction time is less sensitive, than the accuracy rate, to the cognitive load of the primary task when the dual-task condition is operationalized as the particular type of color- or sound-detection task.

Another possible explanation for the lack of simultaneous secondary-task effects for reaction time and accuracy might be the potential competition for attentional resources between reaction time and accuracy. In a dual-task experiment, the primary task is normally given priority and, in this way, the limited residual mental resources that are not being used by the primary task could be allocated to the secondary task (Marcus et al., 1996). For the present study, and the three previous studies mentioned, students have to distinguish whether the sound/colour is correct and, at the same time, react to the right sound/colour quickly when completing the secondary task with limited attentional resources. Completing the secondary task requires participants to shift attention from determining whether the stimulus is right or wrong to producing a response to the stimulus. During this process, students may not equally allocate attention to noticing the correct stimulus (accuracy) and responding to the stimulus (reaction time). According to Wickens' (2007) SEEV model of selective attention, participants' effort of retrieval in the correct sound stimulus might be traded off against the effort of responding to the sound. This suggests that the more effort on reaction time, the fewer differences in reaction time would be found for the simple primary task versus the complex task versions, and accuracy would be curtailed. With the increased cognitive load imposed by the primary task, the mental resources for the secondary task would be more limited and, consequently, the competition might be more intense between the secondary-task reaction time and accuracy. If more attention is directed to reacting to the sound, less attention is allocated to achieving accuracy, and vice versa. Thus, when students complete the complex primary task, two possible result patterns might be found. One is that participants might produce slightly increasing secondary-task reaction time (sometimes reaction time might not be detectable) but statistically rising error rates and decreasing accuracy rates, such as found in Experiment 1 of the present study. Another possible pattern might be, inversely, significantly more reaction time but less detectable changes in accuracy, which was found in Experiment 2 (discussed subsequently). This is probably why when participants completed the complex primary writing task in

Experiment 1, the supposed-to-be longer reaction time was not found to be statistically increased, but the accuracy rate significantly declined in the secondary task.

The findings for two types of error rates seem to respond to the trade-off effects between accuracy and reaction time. Specifically, non-significant differences in secondary-task Error I rate, according to the formulas (see Section 3.5.4.2), indicate that the frequency, with which students missed the correct sound, did not significantly differ between the simple and complex primary writing task conditions. In contrast, the statistically higher Error II rate for the complex primary writing task, than the simple one, suggests that the frequency of participants' reaction to the non-required sound significantly increased. Taken together, the results for Error I and Error II indicate that participants probably sacrificed their attention to distinguish the correct stimulus for speed of responses to the sound.

With dual-task Experiment 2, the reaction time for the secondary task offers evidence in support of the manipulation of two primary writing tasks. Increases in the cognitive-task variability of the primary writing task statistically increased participants' reaction time in the secondary task. No significant decrease, however, was identified for secondary-task accuracy rate when students completed the complex primary writing task versus the simple one. Although patterns for reaction time and accuracy rate, found in Experiment 2, were seemingly opposite to those in Experiment 1, the difference might be due to the designs of the secondary task. Different secondary tasks were connected to varying information processing steps, making it possible to identify on which of the processes the cognitive load is imposed (Baddeley, 1986; Brünken et al., 2003). The secondary task in Experiment 1 was a sound-detection task, in which participants were asked to react fast, and accurately, to a specific sound and ignore another sound. Participants in Experiment 2 were instructed to respond to only one sound, although the stimulus appeared more frequently in comparison with Experiment 1. The designs of the secondary task led to different information processing steps and directed participants' attention to different aspects, reaction time or accuracy. As previously discussed, students appeared to focus on reaction time more than accuracy in Experiment 1, leading to significant differences in the secondary-task accuracy rate. A similar example is the colour-detection task in previous studies (Révész et al., 2016, 2014). Competition between accuracy and reaction time in the secondary task would suggest that participants directed more attention to accuracy than reaction time when the secondary task is operationalized as a tone-response task in Experiment 2 instead of the stimulus-detection task. As allocation of participants' attentional resources during the secondary task has not yet been understood in enough detail, retrospective interviews may be an appropriate way to gain insight into this issue in the future studies.

Differences in findings of Experiment 1 and Experiment 2, to some extent, highlight the importance of the design of the secondary tasks when investigating the cognitive load imposed by primary tasks. To design an appropriate secondary task to balance the potential trade-off effects between reaction time and accuracy, while sensitively reflecting the cognitive load imposed by the primary tasks, is of importance. Some suggestions have been provided to design an appropriate secondary task in the cognitive psychology field (Brünken et al., 2003; Marcus et al., 1996; Schoor et al., 2012). Secondary tasks, however, need to be developed further and designed specifically for the field of SLA, especially the writing field, to explore deep insights into what constitutes an ideal secondary task in the dual-task paradigm in future research.

To conclude, the results obtained from the dual-task experiments, that is, accuracy rates in Experiment 1 and reaction time in Experiment 2, closely align with the findings of the self-ratings on task difficulty and mental effort. These results suggest that these indicators measure the same construct, cognitive load, but from different perspectives (Brünken et al., 2010). Taken together, the findings appear to indicate that the cognitive-task variability manipulations in this research were successful: The complex task version, as intended, yielded greater cognitive demands than the simple version.

6.3 Effects of Cognitive-task Variability and Planning Type on Performance by Groups

This section addresses RQ2: What are the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing? Firstly, the effects of cognitive-task variability on writing performance are interpreted by planning groups. Secondly, for each task versions, the simple task and the complex task, the effects of planning type are compared and interpreted. Finally, the overall effects of cognitive-task variability and planning type on writing performance are summarized and discussed.

6.3.1 Effects by pre-task planning group

Participants in the pre-task planning group were asked to do a 15-minute pre-task planning and 25-minute quick writing. The focus of this discussion is on two major findings for the PTP group. The results for the effects of cognitive-task variability on EFL writing performance in the PTP group are discussed, with the differences between the PTP and the CG in learners' writing performance compared.

6.3.1.1 Effects on adequacy

Adequacy is evaluated by subjective rating of content, organization and overall text quality, based on Jacobs et al.'s (1981) analytical rating scheme, in which content (30%) and organisation (20%) account for 50% of the overall text quality.

With regard to the effects of cognitive-task variability, PTP students in the complex task obtained statistically better performance in content and organisation, and an overall writing quality that was higher, but not significantly, than the simple task. Rahimi and Zhang's (2018) study, similarly reported that increasing cognitive-task variability positively affected learners' performance in terms of content, organisation and overall score under the PTP condition.

The favourable results in students' adequacy performance yielded by increasing the cognitive-task variability and the provision of the pre-task planning are compatible with Robinson's (2001b, 2005) CH. According to Robinson, simultaneously increasing task complexity along the resource-directing dimension, and reducing task complexity along the resource-dispersing dimension, will positively influence students' EFL production. There were no explicit predictions in CH on adequacy, however. When greater resource demands are imposed in the complex task, pre-task planning can help ease participants' procedural demands on the working memory. In this way, participants are likely to reach an optimum resource allocation to satisfy the pragmatic demands of the task, and thus achieve positive outcomes for adequacy.

Comparison of students' adequacy performance between the CG and PTP groups indicated that the CG gained a higher, but not significant, overall score than the PTP group in the simple task. In complex task, the PTP group, however, outperformed the CG in content, organisation and overall writing quality, although only content achieved statistical significance. It seems that pre-task planning is more useful for the task that is more cognitively taxing. A rationale for doing pre-task planning in the complex task is that there is more scope for effective preparation for content and perhaps, organisation, than with the CG condition, thus freeing learners' attentional resources during task completion (Skehan & Foster, 1997), and resulting in better overall text quality. The simple task, combined with pre-task planning, may greatly lessen the challenge of cognitive task complexity. If a task is easy, learners are able to perform well without the need for planning (Ellis, 2005), which is possibly the reason participants' adequacy performance did not benefit from the provision of pre-task planning in the simple task, as compared with the CG condition.

Pre-task planning resulted in better performance with content and organisation in the complex task than the CG condition, consistent with the attentional funnelling model posited by Kellogg et al. (2013). According to Kellogg et al. (2013), prewriting activities, such as topic outline before writing, that is, pre-task planning, can funnel learners' attention to the macrostructure instead of microstructure, allowing more time and effort to generate and organise ideas. Galbraith and Torrance's (2004) study also showed that planning notes in advance aided writing regardless of whether the notes were available for writing. The interview data provide evidence of participants' focus during pre-task planning. PTP interviewees reported that during pre-task planning they focused mainly on content (55.56%) and organisation (33.33%) by using drafting and organisational planning strategies. Some interviewees stated that pre-task planning helped them organize their thoughts and ensured their writing was more topics-related.

Taken together, increasing cognitive-task variability might activate writers' working memory for the formulation process (Kormos, 2011). Pre-task planning appears to facilitate the formulation stage, especially the *planning* sub-process, as it enables writers to plan ideas based on the input provided in the task prompts, or retrieved from long-term memory and organise ideas in a coherent order. The qualitative data suggest that the PTP interviewees focused mainly on the *planning* sub-process of formulation during pre-task planning. With the help of pre-task planning, PTP writers are more likely to perform better than the CG participants in the complex task (rather than the simple task) in terms of content and organisation, and, consequently, overall quality.

6.3.1.2 Effects on fluency

Increasing cognitive-task variability positively affected pre-task planners' fluency, which differs from the results obtained by Rahimi and Zhang (2018). Two possible reasons may account for the divergent findings. One is related to the topic, which in Rahimi and Zhang's (2018) study was how to allocate funds to varying projects, while students, in the present study, were asked to choose their best roommates. In Chinese universities, students live in dormitories and, consequently, they know a lot about living with a roommate in a dormitory and probably have criteria for the best roommates. According to Kellogg (2013), writing a familiar topic "provides a scaffold to support the writer and avoid overloading the limited resources of working memory" (p. 183). Writing about the "roommate" topic that participants know well, therefore, might free their short-term working memory for the task. It would permit more executive attention to be allocated to formulation, *executing* and/or *monitoring* and lead to a faster writing speed than when writing about a less well-known topic. When cognitive-task variability was increased, in the present study, writing a familiar topic may have contributed to the favourable results for the PTP group in fluency. In contrast, when students were asked to complete an unfamiliar topic, as in Rahimi and Zhang's (2018) study, they wrote less fluently in the complex task than the simple task.

Another reason may be the explicit operationalising of pre-task planning. In Rahimi and Zhang's (2018) study, students were given 10 minutes to do the pre-task planning, while in the present study a 15-minute pre-task planning was provided. The extra five-minute planning time might have helped students prepare better, especially for a complex task, thus contributing to greater fluency by generating more ideas and making more organized structures. An insight is provided in one participant's interview data: Participant #08 expressed that she prepared well for writing the two tasks both physically and mentally during the 15-minute pre-task planning, compared to the short pre-task planning she did previously. She found she could think about the structure and content more thoroughly, make a clearer outline and write more fluently than before.

A comparison of the fluency of the PTP and the CG shows that, in the complex task, the PTP group statistically outperformed the CG, while in the simple task their fluency was similar. This pattern, similar to those found in adequacy results, aligns with the results in previous oral studies (Foster & Skehan, 1996; Skehan & Foster, 1997) in which a strong positive influence of planning on fluency was found in the complex task. These consistent findings provide more evidence and support for Ellis' (2005) conclusions that "planning has a greater effect on production in general when the task is cognitively demanding" (p. 24). As discussed in Section 6.3.1.1, decreases in cognitive-task variability along with the provision of pre-task planning may reduce cognitive task complexity. It can be inferred the simple task might be too easy for the PTP participants and that "if tasks are too easy, they will present no challenge, and are not likely to extend any other goals of restructuring, accuracy, or fluency in any effective way" (Skehan, 1996b, p. 53). This might account for the similar fluency performance found between the CG and PTP groups in the simple task.

In contrast, when the task is more cognitively demanding, it seems that pre-task planning facilitates the *planning* process, which helps writers set goals and generate and organize relative ideas for the complex task. Such macro-level preparation is reflected in PTP participants' performance in content and organisation in the complex task (see Section 6.3.1.1), and also in participants' responses in the interview. It can be assumed that writers write more quickly and fluently with a lessened pressure on working memory, if they have had clearly structured ideas as the task required, listed the relevant statements with supporting points and outlined the information in an organized way before writing. Another reason for pre-task planners' writing more fluently might be related to affective factors (Ellis & Yuan, 2004). As evident in the interview, pre-task planning may help reduce participants' writing anxiety, increasing their confidence in completing the task, especially when performing a complex writing task. This probably reduces the number of participants' "revisions undertaken in L2 writing" and lessens "their need to engage in extensive *monitoring*" (Ellis & Yuan, 2004, p. 78).

6.3.1.3 Effects on accuracy

Students under the PTP condition wrote significantly less accurately, compared to those under the CG condition, in both the simple and complex task versions, although increasing cognitive-task variability resulted in a marked rise in accuracy in terms of EP100.

The findings of lower accuracy scores in the simple task than the complex task again confirmed the researcher's speculation that the simple task might be too easy when participants were also provided with pre-task planning. When the task is not challenging, participants are not likely to activate any writing goals of accuracy, complexity or fluency effectively. (Ellis, 2005; Skehan, 1996b). In contrast, increasing cognitive-task variability augmented the difficulty of the task, providing the PTP participants with a greater challenge. The complex task, therefore, may fully trigger PTP participants' central executive in working memory with a consequence fewer errors in their output.

Pre-task planners were pressured to write quickly in 25 minutes so that relatively less working memory was available for the *monitoring* process to edit their internal thinking and external output. In comparison, CG writers, who were asked to write in their own way, may have had more additional time at their disposal during writing to carefully attend to the *monitoring* process in both the simple and complex task versions, which may have led to a greater focus on linguistic accuracy. The interview data provide evidence that participants from the CG focused mainly on monitoring linguistic accuracy during the L2 writing process. It can be inferred that the differences in accuracy performance between the PTP and the CG is due to the extent to which participants' working memory was allocated to the *monitoring* process under the PTP and CG conditions. Furthermore, whereas 3 out of the 5 CG interviewees stated that they had a few minutes left for checking the grammar errors, all pre-task planners in the interview reported they had no time left during the 25-minute writing. The extra "editing" time available to the CG writers may have also contributed to better accuracy in writing of participants in the CG than the PTP group.

The difference in accuracy between the CG and PTP groups is inconsistent with the findings obtained by Ellis and Yuan (2004) and Rahimi and Zhang (2018). As explained by Ellis (2005), the mixed findings in the effects of pre-task planning on accuracy may be due to a variety of factors, such as "the extent to which particular learners are oriented towards accuracy, the learners' level of proficiency, the type of task, the length of planning time available and the particular grammatical feature" (p. 22). Further empirical research is warranted to take these variables into consideration, providing a comprehensive picture of the effects of pre-task planning on accuracy.

6.3.1.4 Effects on syntactic complexity

The effects of cognitive-task variability on syntactic complexity were examined in this study. The four sub-constructs of syntactic complexity were: general index (MLT) to present the overall complexity; coordination (T/S) as an index of complexification at beginning levels; subordination (C/T and DC/C) at intermediate levels, and phrasal-level complexification (MLC, CP/C and CN/C) at advanced levels.

Increases in cognitive-task variability positively affected PTP participants' syntactic complexity only in terms of CP/C while the other six measures of syntactic complexity were not significantly changed. The improvement in CP/C lends tentative support to Yoon's (2017) argument that "phrase-level complexity measures are capable of detecting L2 writing development more sensitively than clause-level measures" (p. 132). Results from the current study agree with those of previous studies (Gilabert, 2007; Rahimi & Zhang, 2018). In Gilabert's (2007) research, a marginally better syntactic complexity performance was found in the complex task than the simple task. Similarly, Rahimi and Zhang (2018) found that increasing cognitive-task variability enhanced participants' syntactic complexity in terms of subordinate clauses under the planning condition. It is interesting to notice that there is an increase in subordination but no significant changes in CP/C in Rahimi and Zhang's research (2018), while an increase in phrasal coordination but no salient variance in subordination were identified in the present study. This seemingly contradictory finding may be due to participants' different proficiency levels, as subordination was a sensitive measurement at intermediate level and phrasal-level elaboration at advanced level in the field of writing (Norris & Ortega, 2009; Yang et al., 2015; Yoon, 2017). The favourable results of syntactic complexity under the PTP condition, however, should be cautiously interpreted (Pallotti, 2009), since only one measure (CP/C) out of the seven in the present study, one index (subordination) out of three in Rahimi and Zhang's (2018) research and no syntactic measures in Gilabert's study (2007) reached significance. Further empirical evidence from a larger body of data is warranted to provide further insights into whether and to what degree increasing task complexity positively influences students' syntactic performance under the PTP condition.

No significant differences were found in syntactic complexity between the CG and PTP groups in either simple or complex task versions, although the CG gained a marginally higher performance in MLT, C/T, DC/C, T/S and CN/C than the PTP group. Taken together, it can be inferred that pre-task planning may not benefit students' syntactic performance. A likely explanation is that, according to Kellogg's (1996) model, students' central executive has limited capacity; as a result, writers probably have to prioritise specific writing process/processes when under pressure to rapidly complete a writing task. Consequently, the writing processes and sub-processes may compete for attentional resources during pre-task planning and writing. To elaborate, during pre-task planning, participants were likely to use more working memory in drawing up a conceptual plan of the task (the *planning* sub-process) than in formulating detailed linguistic plans (the *translation* sub-process). During writing, pre-task planners had to deal with the formulation, execution and *monitoring* processes in a short time (25 minutes). In this way, PTP participants when writing, compared to those under the CG condition, may be pressured to encode the ideas planned during pre-task planning into linguistic frames that are easy to access quickly, with limited attention devoted to complexifying linguistic encoding and control over the existing interlanguage system. A better performance in adequacy and fluency, with similar syntactic-complexity, but less accuracy, under the PTP than the CG condition, seems to support the above predictions. The interview data also provide some evidence that the PTP interviewees preferred more global planning, such as content and organisation rather than local planning, such as linguistic frames, during pre-task planning. Alternatively, ideas and content outlined seem to fade less than the forms planned during pre-task planning (Skehan, 2014a). PTP participants might better retrieve the content and ideas than the syntactic frames they planned during pre-task planning; this may be one of the reasons that the PTP condition, compared to CG, benefited learners' adequacy performance, but marginally influence syntactic performance.

The non-significant differences in syntactic complexity between the PTP and the CG are in contrast to the findings of previous studies (Ellis & Yuan, 2004; Rahimi & Zhang, 2018) in which they found pre-task planning had a greater positive effect on syntactic complexity than in the no-planning (NP) group. The inconsistent results might be due to the distinct implementation of planning conditions. In the prior studies (Ellis & Yuan, 2004; Rahimi & Zhang, 2018), participants spent more time on the tasks in the PTP than the NP group, whereas in the current study the total for the task time for the PTP and the CG is the same. The additional planning time for the PTP group in Ellis and Yuan's

(2004) and Rahimi and Zhang's (2018) studies might help ease writers' working memory and enable them to have more time for the *translating* process (online planning), resulting in better performance in syntactic complexity than for the NP group. In contrast, the present study controlled the total time for the tasks in the PTP and the CG to ensure that any effects of pre-task planning resulted from the given planning time in the formulation process instead of the varying time on the task (Ong & Zhang, 2010).

6.3.1.5 Effects on lexical complexity

Two measures were used to examine the effect of cognitive-task variability and planning type on learners' lexical complexity. Lexical complexity was measured by lexical density and lexical diversity with WT/√2W as a measure.

The results showed that PTP participants' lexical complexity was impervious to the increases in cognitive-task variability due, possibly, to the limited lexical resources of the participants. Some researchers (e.g., Cui, 2011) found that the number of advanced words that can be fully automatized by second-year English majors in China is small. The interview data lend some support that some PTP participants expressed their concern about using simple words and the very limited lexical resources they had controlled. It can be assumed that, in the present study, the increased cognitive-task variability and provision of pre-task planning cannot remove the ceiling imposed by students' limited EFL lexical knowledge. Learners are likely to experience lexical translating problematic issues when they had limited EFL lexical resources or difficulty in accessing such resources.

In the comparisons between the PTP and the CG, participants' lexical performance under the PTP condition was significantly less than those under CG in terms of WT/√2W in the simple task. No significant differences in lexical complexity between the CG and PTP groups were identified in the complex task. This might be because that after the 15-minute pre-task planning, participants were given 25 minutes to write the task with no less than 250 words. In this way, students in the PTP group were supposed to write fast with limited attentional resources devoted to the *translating* sub-process to retrieve advanced or a variety of words during writing. Consequently, they probably chose the familiar or simple words which occurred to them first, so they could complete the writing tasks in a short time. Students in the CG, in contrast, were likely to have more time (especially in the simple task) than those in the PTP group to activate their lexical resources during formulation. As a result, in the simple task the CG performed better in lexical diversity than the PTP group. For the complex task, the CG appeared to lose its advantages in lexical complexity when the increased cognitive-task variability reduced students' attention available devoted to lexical complexity.

The comparison between the PTP group and the CG in the complex task, but not in the simple task, is consistent with the findings obtained by Rahimi and Zhang (2018) who found pre-task planning did not lead to any significant changes in lexical complexity for either the simple or the complex task. As discussed in Section 6.3.1.2, the topic of "choosing the best roommate" in the present study might be more familiar to participants than the topic of "allocating funds to projects" in Rahimi and Zhang's (2018) research, which may reduce the complexity of the task in the present study. In this way, the simple task in the present study might be so easy for the PTP participants that they do not activate their lexical resources (Ellis, 2005; Skehan, 1996b), resulting in lower scores in WT/√2W in the PTP group than the CG. This may explain the difference in findings for the simple task in the present study and Rahimi and Zhang's (2018) research. An alternative reason might be that in contrast to the present study, PTP students in Rahimi and Zhang's (2018) study had enough time to write the essay (10 minutes planning time and 35 minutes writing time) compared with those in the NP (35 minutes writing time). Thus, PTP students could devote as much attentional resources as the NP students to the *translating* process to retrieve their lexical knowledge, leading to similar performance in lexical complexity in the simple task between the PTP and NP groups.

An increase in accuracy and a marginal rise in syntactic complexity were found when cognitive-task variability was increased. The results for the PTP group, when accuracy and complexity (i.e., syntactic and lexical complexity) are taken together, lend partial support to Robinson's CH.

6.3.2 Effects by online planning group

Participants under the OLP condition were asked to start writing immediately to complete two tasks with different levels of cognitive-task variability and to plan during their writing. The focus of this discussion is on two major findings for the OLP group. Firstly, the results for the effects of cognitive-task variability on EFL writing performance in the OLP group are discussed. Secondly, the differences in learners' writing performance between the OLP and PTP groups, then between the OLP group and the CG are compared.

6.3.2.1 Effects on syntactic complexity

The results showed a decrease in C/T and DC/C, and an increase in CP/C under the OLP condition. Ostensibly, the decline in subordination in terms of C/T and DC/C is contrary to Robinson's CH, since increasing cognitive-task variability negatively affected the number of students' subordinate clauses. However, it should be taken into account that learners develop their ability to use phrasal-level complexification later than subordination (Halliday & Mattiesen, 1999). The increases in phrasal-level elaboration (i.e., CP/C) but decreases in subordination (i.e., C/T and DC/C) indicate that online planners in the complex task produced more advanced language, with "lower levels of subordination" but "more complex phrases" (Norris & Ortega, 2009, pp. 562-563) than in the simple task. The favourable results of syntactic complexity under the OLP condition in the complex task, essentially, echo Robinson's CH (2001b, 2005, 2007a). As assumed by Robinson, learners' attention and effort are expected to be channelled to the language code system to meet the increasing conceptual demands, consequently facilitating the interlanguage development, when cognitive-task variability is augmented with more elements and more reasoning demands involved.

Furthermore, the findings of syntactic complexity in the OLP group corroborate the results for phrase-level complexity reported in the previous studies (Biber et al., 2011; Bulté & Housen, 2014; Lu, 2011; Mazgutova & Kormos, 2015; Yoon, 2017), providing further evidence of the validity of phrasal-level elaboration as a sensitive development indicator in EFL writing. The results in the present study also confirm the necessity to use syntactic constructs in different dimensions to capture a multidimensional and dynamic picture of L2 learners' writing development (Norris & Ortega, 2009; Yang et al., 2015; Yoon, 2017).

When compared to the PTP group, in the simple task, the OLP group outperformed the PTP group in DC/C, and in the complex task, the OLP group similarly performed better syntactically, but not significantly different, in terms of C/T, DC/C and CN/C. The higher scores may be because participants under the OLP condition are likely to allocate more attentional resources to the *translating* sub-process than those under the PTP condition. The OLP participants had enough time (40 minutes) to write and do online planning, while their counterparts under the PTP condition had to write fast (25 minutes) with limited time for online planning during writing. OLP participants are likely to take advantage of the additional time available to select complex syntactic frames to encode their ideas (e.g., produce dependent clauses to meet the reasoning demands). According to some researchers (e.g., Ellis, 2005), students did more micro-planning than macro-level planning during online planning. In the interviews, OLP participants reported focusing mainly on language form (61.54%) during writing and online planning; two interviewees stated explicitly that they tried to search for complex sentence patterns when doing online planning.

The comparisons of syntactic complexity between the OLP and PTP groups are contrary to the findings obtained by Ellis and Yuan (2004) who found the pre-task planners used a greater variety of verb forms (an index of syntactic variety). The divergent results might be due to the difference in total time used for completing the tasks in these two studies. In contrast to the present study in which the total time among the three groups was controlled, participants in the PTP group in Ellis and Yuan's (2004) study spent more time (27 minutes) on the task than those in the OLP group (21 minutes). As a result, PTP participants in Ellis & Yuan's (2004) research, as they reported in the post-task interview, have had additional time for the online process and retrieving complex verb forms during the *translating* process.

When compared to the CG, similar syntactic complexity performance was found between the OLP group and the CG in either simple or complex task versions. It can be assumed from the results that both the OLP and CG conditions could enable attentional resources to be paid to the *translating* process and help ease participants' online processing burden, as long as writers have enough time to complete the tasks during the writing process.

To conclude, it is plausible that students, with the help of online planning and increased cognitive-task variability, might be encouraged to use new structures, or to extend their interlanguage system, to the next level of syntactic development in L2 writing to express their meanings.

6.3.2.2 Effects on lexical complexity

Increasing cognitive-task variability reduced OLP learners' lexical complexity production in terms of WT/√2W. This might be because the increased cognitive demands and task difficulty augmented participants' pressure on working memory to complete the task. As speculated in Section 6.3.2.1, online planners are likely to prioritize their attention to the *translating* process of formulation. Cognitive task complexity is supposed to be inherent in the formulation stage of L2 writing process (Kormos, 2011), therefore competition for working memory within the *translating* process (i.e., between selecting lexical units and building syntactic frames) may occur when cognitive-task variability is increased. In this way, less attentional resources would be available for online planners to retrieve advanced lexical units with more working memory allocated to the syntactic structures, when they completed the complex task. As a result, they were more likely to employ familiar or simple words to complete the complex task, compared to the simple task, resulting in a decline in lexical diversity but an increase in syntactic complexity in the complex task. This study extends Frear and Bitchener's (2015) hypothesis that there may be a trade-off between syntactic and lexical means of expression when learners experience pressure which is "brought to bear on limited attentional resources by cognitive task complexity" and pressure "of producing language that was not fully automatized" (p. 53).

When compared to the PTP, the OLP group yielded significantly more lexical complexity in the simple task, and insignificantly better lexical performance in the complex task than the PTP group in terms of WT/√2W. A feasible explanation would be that pre-task planners were more pressured to complete the task faster than the online planners who had enough time to write. The results suggest that online planners used their additional time at their disposal to deal with the *translating* process to retrieve more diverse lexical units than the pre-task planners, especially for the simple task. Whereas, PTP participants' fast writing is less likely to facilitate students' lexical encoding during the *translating* process. Alternatively, it is possible that even when pre-task planners made some lexical preparation during pre-task planning, they may find it difficult to carry over the lexical units they have planned into the textual performance when they are pressured to write fast (Skehan, 2014a).

When compared to the CG, the OLP group significantly outperformed the CG in terms of WT/√2W in the complex task, but the difference in lexical performance between the OLP and the CG did not reach significance in the simple task. The results indicate that in the simple task, participants under the OLP and CG conditions have enough attentional resources allocated to *translating* process, which leads to similar lexical (and syntactic) performance. Increasing cognitive-task variability might push students to make decisions about which writing process to attend to, primarily, during writing. In such situations, CG students might devote less attention, when completing the complex task, to retrieving the lexical storage in their long-term working memory than those under the OLP condition. As a consequence, participants in the OLP group produce better lexical performance than the CG in the complex task.

6.3.2.3 Effects on accuracy

No significant changes in any measures of accuracy were found for the OLP group between the simple and complex tasks. The results indicate that OLP students devoted similar attentional resources to monitor the linguistic accuracy when completing the simple and complex task versions. This finding is contrary to Skehan's LACM, as no trade-off between accuracy and complexity (either syntactic or lexical complexity) under the OLP condition was identified.

When compared to the PTP group, the online planners wrote more accurately than the pre-task planners in both simple and complex tasks, although these differences were not statistically significant. This result suggests that participants under the OLP condition paid marginally more attention to the *monitoring* process than those under the PTP condition who did a fast writing. This finding is congruent with Ellis and Yuan's (2004) comparison between the OLP and PTP groups, although the design for the implementation of planning and task type in the two studies was differently.

Compared to the CG, participants in the OLP group made significantly more Err/T and EP100 in the complex task and performed less accurately, but insignificantly, in the simple task than those in the CG. Online planners are likely to engage in translating the ideas at the expense of monitoring their output when cognitive-task variability was increased. Whereas, students in the CG may take advantage of time available to monitor their internally processed output during *translating* before *executing*, or to edit and correct the external output after they execute the text. The results that the CG condition has a markedly positive effect on linguistic accuracy, compared to the OLP, especially in the complex task, suggest that students, when writing in their own way, prioritize their working memory on monitoring their output accuracy. The interview data for the CG indicate that for some CG interviewees accuracy came first, and they made great efforts to express their ideas accurately. Some of them also reported that they spent any minutes left on checking grammar errors. Students' focus on accuracy might be because teachers are likely to "use grammatical accuracy as the main index of language proficiency" (Trebits, 2016, p. 172) for exams, and highlight students' grammatical errors in their written work when giving feedback (Kuiken & Vedder, 2019; Lee, 2008, 2009, 2011; Mao & Crosthwaite, 2019). These teachers' practices might encourage students to pay attention to their linguistic accuracy when they write.

6.3.2.4 Effects on adequacy

Increasing cognitive-task variability resulted in a significant decrease in OLP participants' adequacy performance in terms of content, organisation and overall writing quality. Increased cognitive-task variability may have greatly augmented OLP learners' pressure on working memory which has a limited capacity (Skehan, 1996b). Participants under the OLP condition were probably pressured to focus more on form at the expense of achieving semantic and pragmatic goals when the writing task was cognitively taxing.

There appears to be a trade-off between syntactic complexity and adequacy performance in the OLP group. Some scholars have also suggested trade-off effects between learners' linguistic outcomes and higher-order dimensions of EFL production. For instance, Pallotti (2009) posited that the structures produced by students may be syntactically complex or accurate but pragmatically inadequate. Similarly, Kuiken and Vedder (2008) expressed concerns that learners' favourable linguistic outcomes might be achieved at the cost of attention being taken away from the higher-order language processes. Kormos (2011) also argued L2 writers' attentional resources might be traded off between syntactic encoding and global organisation of the text during the writing process, and between linguistic accuracy and discourse structure in the *monitoring* process. Kellogg's (1996) writing model seems to provide some insights into such trade-off effect. It can be inferred that the trade-off between learners' linguistic output and adequacy may result from the competition for working memory between the sub-processes of formulation (i.e., between *planning* and *translating*) under certain circumstances (e.g. the OLP condition). This competition may become increasingly obvious when cognitive-task variability is augmented, as reflected in the responses in the interview of the OLP group. Online planners paid the most attention (61.54%) to linguistic production during writing, with limited attention devoted to content (30.77%) and organisation (7.79%).

A significant correlation was found between WT/√2W and the overall writing score in the present study (ρ=.346, p<.001). This result corroborates Lu's (2012) findings, when analysing large-scale date from a corpus of Chinese learners, that Chinese students' lexical diversity is significantly correlated to the raters' judgement of the task quality. The decline in overall quality in the complex task, in the present study, might be related to the decreasing lexical diversity under the OLP condition.

To conclude, it appears that the effects of cognitive-task variability on adequacy as well as syntactic complexity in the OLP group lend support to the LACM, that learners' attentional resources are limited in capacity, although LACM does not make explicit predictions concerning the effect of task complexity on adequacy.

The OLP group was statistically outperformed by the PTP group in terms of content and organisation in the complex task, whereas no significant differences in content, organisation and overall quality were found between the two groups in the simple task. The attention that online planners gave to *translating*, may explain their inability to engage in planning the content and organisation, when the cognitive-task variability was increased. Apart from the *translating* process, OLP participants' better performance in accuracy than the PTP group is evidence that online planners probably focus on the linguistic output rather than the macro-level adequacy during the *monitoring* process. In contrast, PTP, as discussed in Section 6.3.1.1, assists students' output of content and organisation planned during *planning*, thus reducing the pressure on the central executive in working memory during writing, when the level of cognitive-task variability is high. In this way, pre-task planning probably facilitates the achievement of pragmatic goals of content and organisation in the complex task. Online planners might not experience great pressure on working memory when completing the simple task, which helps attentional resources be able to be directed to the *planning* process to generate ideas, leading to a satisfactory performance in content and organisation in the simple task. Alternatively, it is possible that the simple task, as discussed previously (see Section 6.3.1.2), is too easy for the participants, when pre-task planning is provided, leading to poor performance in adequacy as well as in accuracy and syntactic complexity.

Compared to the CG, no significant differences in adequacy were identified between the OLP and the CG in either simple task or complex task. It can be assumed from the results that both the OLP and CG writers allocated similar amount of attention to the adequacy performance.

6.3.2.5 Effects on fluency

Increases in cognitive-task variability did not yield any significant differences in OLP students' fluency performance. According to Skehan's LACM, learners have limited attentional capacity when performing cognitively demanding tasks. Students in the OLP group are likely to prioritize restructuring, which requires attention being devoted to the new forms of language, with accuracy and fluency being secondary, and adequacy possibly the last. This may explain why increasing the complexity of a task has not led to any significant changes in learners' accuracy or fluency, but a significant decline in adequacy in the OLP group.

No significant differences in fluency were identified between the OLP and PTP groups in either simple or complex task versions, but the PTP wrote marginally more fluently than the OLP in the complex task. This might be because with the provision of pre-task planning, students found it slightly easier, causing less anxiety, to access the interlanguage system than students under the OLP condition, thus promoting more fluent writing when the task is cognitively demanding.

The OLP group significantly outperformed the CG in fluency in the complex task, while no statistical significance in fluency was identified between the OLP and the CG in the simple task. A possible reason is that participants under the OLP and CG conditions channelled their limited working memory to different writing processes when completing the complex task. Online planners are more likely, as discussed in Section 6.3.2.1 and 6.3.2.2, to carefully translate the ideas using a range of vocabulary and complex syntactic frames, which may encourage them to write a long passage. Conversely, students in the CG tend to use the additional time at their disposal to carefully attend to linguistic accuracy, which easily inhibits their production of a long written text and reduces their fluency.

From the comparison among the OLP, PTP and CG conditions, it seems that planning (i.e., PTP and OLP) encourages students to write more quickly, or a longer essay, than those under the CG condition, which promotes their writing fluency in the complex task. It might also be inferred that devoting attention to the formulation process (either *planning* or *translating*) facilitates participants' fluency rather than their focus on the *monitoring* process when the writing task is cognitively taxing.

Findings from the comparison among the three groups were contrary to the results obtained by Ellis and Yuan (2004) who found that pre-task planners had a definite advantage in fluency over those in the OLP and NP groups, whose fluency performance was similar. The divergent results may be due to two underlying reasons. One is related to the different task types, as different impacts of task types have been found on L2 production (Skehan, 2009). Ellis and Yuan (2004) used the narrative writing task, without task complexity manipulation, while the present study used two argumentative writing tasks with cognitive-task variability manipulated by varying numbers of elements and reasoning demands. Another reason may be the differences in implementations of planning in these two studies. Participants in the PTP group in Ellis and Yuan's (2004) study spent most time on the task, followed by those in the OLP group and then the NP group, while the present study controlled the total time among the PTP, OLP and CG conditions. With different times allocated in the planning groups, participants might prioritize their attention differently to the dimensions and processes of writing, leading, therefore, to varying performance in fluency. It is noticeable that online planners in Ellis and Yuan's (2004) study were asked to write without a time limit, which may discourage students to write quickly, thus reducing online planners' fluency, in comparison even with those under NP. Whereas, in the present study, although OLP participants had enough time to complete the task, there was a 40 minutes time set, which may to some extent encourage students to write quickly.

6.3.3 Effects by comparison group

Participants under the CG condition, as in the most previous studies, were asked to complete the tasks in a natural way within 40 minutes (Cho, 2015; Frear & Bitchener, 2015; Kormos, 2011; Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2007b, 2008, 2012; Rahimi, 2019). No pre-task or post-task activities, such as pre-task planning or editing, were involved in the comparison condition. This section interprets the results for the effects of cognitive task complexity on L2 writing performance in the CG in relation to previous empirical studies and theoretical frameworks.

6.3.3.1 Effects on syntactic complexity

Hardly any significant differences in syntactic complexity were found between the simple and complex tasks in the comparison group with no evidence in favour of either the LACM or the CH.

While these findings are in line with some prior research (Cho, 2015; Frear & Bitchener, 2015; Kormos, 2011; Kuiken et al., 2005; Kuiken & Vedder, 2007b, 2008, 2012), they are inconsistent to those of other previous studies (Ishikawa, 2007; Rahimi, 2019; Rahimi & Zhang, 2019). A plausible reason for the divergent results obtained between the present study and those of Ishikawa (2007) is the different research design. In Ishikawa's (2007) study, high-school participants were asked to write two narrative story-telling tasks with cognitive-task variability of the writing tasks manipulated as [± Here-and-Now]. Under the [+Here-and-Now] simple condition, students were allowed to view a strip cartoon while writing and were required to use the present tense when telling the story. Whereas, the strip was removed before writing and the usage of past tense was required in the [-Here-and-Now] complex condition. In the present study, tertiary-level students completed two argumentative writing tasks with varying cognitive demands to do with the numbers of elements and reasoning. Different manipulations of cognitive-task variability may lead to a range of task demands. Participants with varying proficiency levels may have dissimilar perceptions of the task demands, which are then reflected in their effort expended in the tasks. They, thus, could allocate their attention to relevant aspects at the conceptualisation stage, leading, finally, to different writing performance (Robinson, 2011a, 2011b).

The choice of topics may account for the mixed results between Rahimi et al.'s studies (Rahimi, 2019; Rahimi & Zhang, 2019) and the present study, since significant topic effects on L2 argumentative writing performance have been reported in prior studies (Yang et al., 2015; Yoon, 2017). Rahimi et al.'s topic was allocating funds for public projects while the present study used “choosing a roommate” as the topic. These two topics may have elicited different levels of cognitive-task variability, such as a range of reasoning demands, which may influence writing performance.

6.3.3.2 Effects on lexical complexity

Increasing cognitive-task variability resulted in a decrease in CG participants’ lexical performance in terms of LD and WT/√2W. While this tendency did not differ from the results for the OLP group in the present study, it is contrary to some previous studies (Frear & Bitchener, 2015; Kormos, 2011; Kuiken & Vedder, 2007a; Rahimi, 2019); comparison with these studies, however, may be inappropriate due to the divergence in task types, task manipulation, and measures employed in these studies (see Literature Review in Section 2.4). In addition, students’ allocation of attentional resources may account for the mixed results of lexical complexity. For example, as Frear and Bitchener (2015) inferred, writers in their study may have directed more attentional resources to lexical variety at the expense of syntactic complexity, which may have led to an increase in lexical complexity but a decrease in syntactic complexity as a function of raising cognitive-task variability. Participants in the present study, however, may have paid more attention to accuracy (discussed subsequently in Section 6.3.3.3) rather than lexical complexity, which resulted in the decreasing number of lexical words and limited range of vocabulary used in the complex task.

Further empirical research is, therefore, warranted to control for variables, such as task type and task manipulations, and adopt a multidimensional measure of lexical complexity, including “lexical density, lexical diversity, lexical sophistication, and number of errors in vocabulary use” (Read, 2000, p. 191) for a more comprehensive picture for learners’ lexical complexity.

6.3.3.3 Effects on accuracy

The results showed that participants made more Err/T in the simple task than the complex task with no significant changes in EP100 and EFC/C between the simple and complex tasks. These findings agree with the prior studies (Ishikawa, 2007; Kuiken et al., 2005; Kuiken & Vedder, 2007a, 2008) but are contrary to Rahimi’s (2019) study. The divergence of the results may be due to the different topics, as discussed in Section 6.3.1.2. The familiar topic might have lessened the pressure on participants’ limited working memory (Kellogg et al., 2013), resulting better accuracy performance in the complex task in the present study.

Another feasible reason might be related to participants' attention allocation. As discussed in Section 6.3.1.3 and 6.3.2.3, CG students are likely to prioritize their working memory to monitor the linguistic accuracy when writing. Increases in cognitive-task variability have directed writers' attention to the *monitoring* process, resulting in better accuracy performance in the complex task, as evident reflected in the interview data of the comparison group. In contrast, participants' attention in Rahimi's (2019) study may be allocated primarily to *translating* rather than *monitoring* process, because an increase in complexity but a decrease in accuracy was found as a function of raising cognitive-task variability.

In the CG, it appears that there is competition for attention between complexity and accuracy, since the rise in accuracy is accompanied by a decline in lexical complexity when cognitive-task variability is increased. These findings support Skehan's LACM.

6.3.3.4 Effects on adequacy

Increasing cognitive-task variability reduced CG participants' adequacy in terms of content, organisation and overall writing quality, similar to the OLP group in the present study. The results indicate that CG participants did not allocate sufficient attention to adequacy, as processing resources seem to be committed to achieve linguistic accuracy, when there is pressure on the working memory in the complex task. This finding differs with the results obtained by Rahimi (2019), due, possibly, to task topic, research context and participants' proficiency level. The effects of cognitive-task variability on adequacy deserve to be studied further because of the inconsistency in results among the limited number of studies (Rahimi, 2019).

The favourable accuracy performance in this study seems to have been achieved at the cost of adequacy. Likewise, Abrams & Byrd (2016) reported a trade-off between text content and grammatical and lexical accuracy. A number of other studies have also posited that competition for attention between linguistic level and the higher-order language process may exist (Kormos, 2011; Kuiken & Vedder, 2008; Pallotti, 2009). Such interpretations, however, should be treated cautiously, because limited research has examined the effects of cognitive task complexity synchronously on linguistic outcomes and higher-order adequacy.

6.3.3.5 Effects on fluency

Increasing cognitive-task variability did not significantly influence writers' fluency in the CG. This finding is neither congruent with the previous studies where an increase in fluency was found in the complex task (Cho, 2015; Ishikawa, 2007), nor with the studies where a decrease in fluency was detected in the complex task (Ruiz-Funes, 2015). The mixed results in fluency in these studies might be due to diverse proficiency levels of participants, different operationalisations of fluency and varying manipulations of task complexity. For instance, high-school students were recruited in Cho's (2015) and Ishikawa's (2007) studies, whereas participants in the present study, and Ruiz-Funes' (2015) research, were undergraduates. Moreover, measures for fluency in Ishikawa's (2007) and Cho's (2015) research were operationalized as words per T-unit, words per clause and/or words per sentence, while in the present study and Ruiz-Funes' research (2015), words per minute was the measure of fluency.

Task type may also have affected fluency. In the present study and Cho's research (2015), cognitive-task variability within two argumentative writing tasks was manipulated by varying the number of argument elements and reasoning demands. In contrast, Ishikawa (2007) used the cartoon story telling narrative tasks manipulated by [$\pm$Here-and-Now], and in Ruiz-Funes' (2015) research, task complexity was determined by familiarity of topic, genre, task type and reasoning demands.

6.3.4 Overall effects of cognitive-task variability and planning type

This section interprets the results from a more macro-perspective to provide an overall view of the effects of cognitive-task variability and planning type in relation to the theoretical perspectives and frameworks. First, the result patterns for effects of increasing cognitive task complexity in each planning group are summarised and interpreted. Then, the issue of how different planning types may influence different writing process components and interact with the cognitive-task variability is summarised and discussed.

Table 6.1 Result Patterns for the Effects of Increasing Cognitive-task Variability on EFL Writing Performance in the Three Planning Groups

	Syntactic	Accuracy	Lexical	Adequacy	Fluency
PTP	+-	+	=	+	+
OLP	+	=	-	-	=
CG	=	+	-	-	=

Note. + means that increased cognitive-task variability had a positive effect, = means that increased cognitive-task variability had a non-significant effect, - means that increased cognitive-task variability had a negative effect, +- means increased cognitive-task variability only had a positive effect on one out of seven measures, which needs further confirmation.

As shown in Table 6.1, increasing cognitive-task variability manipulated by adding numbers of elements and reasoning demands elicited different result patterns in the three planning groups. With the PTP group, increases in cognitive-task variability resulted in the enhancements in accuracy, fluency, adequacy and perhaps, syntactic complexity. These findings lend support to Robinson's CH that simultaneously increasing cognitive task complexity along resource-directing dimensions and reducing the complexity of resource-dispersing task characteristics will positively affect learners' EFL writing.

With the OLP group, increasing cognitive-task variability led to a trade-off between syntactic and lexical complexity, and between syntactic complexity and adequacy. No trade-off effect between complexity (either syntactic complexity or lexical complexity) and accuracy was found. The results indicate that tensions may exist between syntactic and lexical complexity, and between the local linguistic output (i.e., syntactic complexity) and the higher-order process (i.e., adequacy). These results partially support Skehan's (2013) LACM that "humans have a limited information processing capacity and must therefore prioritise where they allocate their attention" (p. 189).

As for the CG, trade-off effects were identified between lexical complexity and accuracy, and between accuracy and adequacy when cognitive-task variability was increased, although no competition among syntactic complexity, accuracy and fluency was detected. These results suggest that there is competition between local linguistic output (i.e., accuracy) and global performance (i.e., adequacy), which partially reflects Skehan's (2013) LACM.

Table 6.2 Result Patterns for the Effects of Planning Type and Interaction between Cognitive-task Variability and Planning on EFL Writing Performance

	Simple task	Complex task	Interaction: Complexity*planning
Syntactic	OLP>PTP	No Sig.	No Sig.
Lexical	OLP>PTP, CG>PTP	OLP>CG	Sig.
Accuracy	CG>PTP	CG>PTP, CG>OLP	No Sig.
Fluency	No Sig.	PTP>CG, OLP>CG	Sig.
Adequacy	No Sig.	(PTP>CG)[24], PTP>OLP	Sig.

Note. No Sig. indicates no significant effects, Sig. indicates significant interaction effects between cognitive task complexity and planning type.

As summarized in Table 6.2, online planning facilitates writers' syntactic and lexical complexity, particularly with the simple task. The CG condition allows writers to pay more attention to controlling the existing interlanguage system, although it may inhibit their writing speed to some extent. Pre-task planning scaffolds students to focus more on adequacy, especially when the task is cognitively demanding. The results suggest that different planning types reduce pressure on the central executive in working memory in different ways, leading to different effects on learners' L2 production, thus confirming Ellis and Yuan's (2004) conclusion. It seems that online and pre-task planning, compared to CG, facilitate the formulation process but in different ways. Pre-task planning may facilitate the *planning* process, while online planning may primarily assist the *translating* process. The results for the CG indicate that learners seem to allocate more attentional resources to *monitoring* when writing in their own style, which may be influenced by the frequently received local feedback (Kuiken & Vedder, 2019; Lee, 2008, 2009, 2011; Mao & Crosthwaite, 2019) and teachers' emphasis on grammatical accuracy (Trebits, 2016). The OLP group might also allocate some working memory to the *monitoring* process, but pre-task planners are likely to have limited amount of time devoted to *monitoring* compared to the OLP and the CG in the present study.

[24] PTP>CG was found in the results for content, not for organisation and overall score.

Cognitive-task variability interacted with different planning types as shown in Table 6.2, resulting in differential influence on EFL writing production. There was a greater online planning effect on lexical diversity for the students with the simple task rather than the complex one, whereas a greater pre-task planning effect was identified on fluency, content and organisation performance when students completed the complex task. Furthermore, the effects of cognitive-task variability on writing quality differed according to the planning types. The three planning groups got a similar overall writing score in the simple task, while the PTP group surpassed the OLP group and the CG with a higher, but not significantly different, writing quality in the complex task.

To conclude, tasks that were less cognitively demanding when planned online, seem to favour learners' writing performance. Whereas, when tasks are more cognitively demanding, pre-task planning may facilitate writers' production. Additionally, learners' accuracy seems to benefit most from "a natural writing way", although the fluency was inhibited in the complex task. It is possible that a writing pattern with both pre-task planning and online planning may help writers to optimally allocate attentional resources and achieve a balanced writing performance in terms of complexity, accuracy, fluency and adequacy.

6.4 Relationships among Anxiety, Cognitive-task Variability, Planning Type and EFL Writing Performance

The discussion in this section addresses the overarching Research Question 3: What are the relationships among students' perceived anxiety, cognitive-task variability, planning type and EFL writing production? In this section, the quantitative results for RQ3 are discussed and interpreted integrating the qualitative interviews as evidence regarding how individual participants undertook the planning and the tasks.

6.4.1 Cognitive-task variability and writing anxiety

The effects of cognitive-task variability on performance differ for high- and low-anxiety groups. With the high-anxiety group, increasing cognitive-task variability led to a significant decrease in writing performance: C/T, DC/C, WT/√2W under the OLP condition; MLT, DC/C, and CN/C under the PTP condition; and LD, WT/√2W and WQ under the CG condition. In contrast, with the low-anxiety group, students' writing performance was significantly enhanced, as tasks increased in complexity: CP/C under

OLP, content and fluency under PTP, and accuracy (fewer Err/T) under CG. A possible explanation for the divergent trends is that participants with different levels of anxiety benefit differentially from the increases in cognitive-task variability. According to the interview data, 62.5% participants in the high anxiety group, regardless of planning conditions, reported that the complex task made them more nervous than the simple version. Their worries related to task difficulty: How to well organise the essay; how to balance an argumentation; and how to make decisions in a short time. These worries and a heightened state of anxiety may impede cognitive processing by reducing the attention available to the complex task, negatively affecting L2 writing performance (Derakshan & Eysenck, 2009; MacIntyre & Gardner, 1994; Trebits, 2016), and block participants' capacity to efficiently write in L2 in the complex task. In contrast, a small number of low-anxiety participants (43%) reported that they felt more anxious in the complex task than the simple version, and the anxious feelings mainly resulted from task difficulty. In addition, 2 out of 7 low-anxiety interviewees stated that they did not feel obviously nervous for either simple or complex task. Compared to the high-anxiety counterparts, low-anxiety participants were less likely to be affected by the anxiety induced by increased cognitive-task variability. For this reason, the low-anxiety participants may benefit more from increased cognitive-task variability with their attentional resources directed to certain aspects of writing performance; for example, improving in CP/C under the OLP condition, content and fluency under PTP, and accuracy under CG in the complex task. The contrasting findings for the high- and low-anxiety groups, in relation to the effects of cognitive-task variability on writing performance, confirm Robinson's (2011a) CH that learners' affective factors contribute to perceptions of task difficulty, and more clearly differentiate learners' performance in the complex tasks.

High-anxiety pre-task planners' accuracy significantly improved as a function of raising cognitive-task variability. A trade-off effect between accuracy and syntactic complex was identified for the high-anxiety participants under the PTP condition, when the reducing MLT, DC/C, and CN/C production in the complex task were taken into consideration. Enhancement in accuracy at the cost of syntactic complexity might occur because the high-anxiety students often suffer "reduced risk-taking ability" (Oxford, 2000 p. 63), and "learners who dislike risk-taking will, presumably, be drawn to accuracy, because of a reluctance to use language they are not sure of" (Skehan, 1996b,

p. 47). The high-anxiety PTP participants may, therefore, prefer a conservative rather than risk-taking strategy within the 25-minute writing time, promoting accuracy at the expense of syntactic complexity. A similar trade-off between accuracy and complexity in anxious students' performance was reported in the Trebits' (2016) research. This competition between complexity and accuracy confirms Trebits' (2016) predictions that anxiety-prone individuals are less likely to attend simultaneously to both accuracy and complexity, as they experience more difficulties in balancing their writing with different aspects of performance competing for attention.

The positive effect of the increased cognitive-task variability on fluency under the PTP condition, regardless of the anxiety level, is consistent with the results for RQ2 in Section 4.3.4. This favourable impact may be the result of pre-task planning. Participants with high- and low-anxiety, as identified in the interview, agreed that pre-task planning had improved their writing fluency. As discussed earlier in Section 6.3.1.2, students are more likely to write fluently when they have made a good pre-task preparation before writing. In the interview, students observed that pre-task planning reduced their anxiety and increased their confidence in completing the task. This may explain why pre-task planners' writing fluency was not reduced but improved, when the complex task increasingly required students' attention.

6.4.2 Writing anxiety and performance

The low-anxiety pre-task planners had better performance in WT/√2W and overall writing quality for the simple task, and in MLT, WT/√2W, content and overall writing quality for the complex task, than their high-anxiety counterparts. This finding is consistent with previous studies in the field of SLA where writing anxiety has been found to negatively correlate with students' writing performance (Guo & Qin, 2010; Horwitz, 2001; Rahimi & Zhang, 2019); similarly, in the area of cognitive psychology, Derakshan and Eysenck (2009) found "anxiety is associated with adverse effects on cognitive performance" (p. 168). According to Krashen (1982, 1985), a high level of anxiety easily arouses participants' affective filters, inhibiting learners' performance; this explains the better performance of students in the PTP-L group than that in the PTP-H condition.

In the interviews, high-anxiety pre-task planners were likely to use local planning strategies, such as drafting, and L1 during pre-task planning, while their low-anxiety counterparts did more global planning, such as making an outline in L2 and listing a few details they could use. The macro-level planning, such as ideas, structure and organisation seem to fade less than the micro-level one, such as drafting and specific language expressions (Skehan, 2014a). It might be, therefore, assumed that the global planning (i.e., organisational planning) can be better endured and more effectively transferred during writing than the local planning (i.e., drafting). In this way, the low-anxiety pre-task planners may more easily retrieve what they had planned than their high-anxiety counterparts. Some low-anxiety PTP interviewees stated that they could spare extra attention during writing to add details into the essays to support statements, or introduce some connective words to enhance the coherence of the essays. By contrast, the high-anxiety writers seem to be fully occupied in transferring what they have planned, and translating their ideas into L2 with simple language that are easy to access quickly, which, in turn, intensifies their anxiety.

Of note is that high-anxiety students' beliefs of pre-task planning seem contrary to their writing performance. As shown in the interview, 100% of high-anxiety students reported that pre-task planning had reduced their anxious feelings, and their perceptions of pre-task planning were more favourable than those of their low-anxiety counterparts. Based on the interview, high-anxiety participants can be expected to improve their writing performance with the help of pre-task planning, closing the gaps with their low-anxiety counterparts. The differences in writing performance between high- and low-anxiety students under the PTP condition are, however, not reduced as expected. This may be because the diminished anxiety mentioned by the high-anxiety interviewees is the Input and Processing anxiety (Tobias, 1986) rather than the Output anxiety. The pre-task planning may allow high-anxiety students more time to encode external instructions, plan the message and make preparations for the organisation and content, easing their pressure at the Input and Processing stages. The high- and low-anxiety students, however, probably experience quite different anxiety levels at the Output Stage. Evidence can be found in the interview data that some high-anxiety interviewees experienced increased pressure when transferring the plan into on-line performance. Different planning strategies, varying quality of pre-task planning and transfer experience under limited writing time, may result in different levels of Output anxiety for the PTP-H and the PTP-L groups.

The relationship between anxiety and writing performance under the PTP condition is non-linear. On the one hand, divergent planning strategies adopted by students with different levels of anxiety may influence their planning quality, which further impacts the usage of planning during writing and anxious feelings at different stages. On the other hand, students' varying anxious feelings during writing may impact on their retrieval and transfer of the plan into on-line performance, as "less pressuring performance conditions create a better context to remember what has been planned" (Skehan, 2014a, p. 220). Consequently, the low-anxiety pre-task planners performed significantly better in L2 writing than their high-anxiety counterparts.

Notably, the superiority of low-anxiety individuals in writing performance is more obvious in the complex task than the simple version, which again supports Robinson's (2011a) CH that individual differences more clearly differentiate learners' performance in the complex tasks. This result is also consistent with the findings in the field of cognitive psychology that "anxiety impairs performance, especially when the task being performed is complex and attentionally demanding" (Derakshan & Eysenck, 2009, p. 168).

Unexpectedly, no significant differences in writing performance between high- and low-anxiety groups were found for the OLP condition, compared to the PTP condition. Some evidence found in the interview data may explain such phenomenon. Online planners were asked to start writing immediately, which increased their anxious feelings at the Input and Processing stages. The heightened state of anxiety, cumulating at the Input and Processing stages may raise participants' affective filter and slow down the retrieval for the information formed at the prior stages, increasingly inhibiting learners' performance at the Output stage (MacIntyre & Gardner, 1994; Tobias, 1986; Trebits, 2016). In other words, the OLP condition, unlike PTP, may arouse students' writing anxiety at all the three stages during writing. This explains why, not only all the high-anxiety interviewees, but also 1 out of 2 low-anxiety OLP interviewees admitted that their writing anxiety increased when only online planning was involved. In this situation, both high- and low-anxiety students' performance is probably impaired by a heightened state of anxiety under the OLP condition. This may be one of the reasons that there was little difference in performance between low- and high- anxiety students under the OLP condition. Such interpretation, however, should be treated cautiously, due to the small interview sample. Further studies are warranted to confirm this finding. For example, a quantitative approach with open-ended questionnaires could be administered to a large data sample for content analysis; for the qualitative approach, think-aloud methods could be used to gain insights into students' anxiety at the Input, Processing and Output stages during the online planning writing process.

6.4.3 Planning conditions and writing anxiety

For the high-anxiety group, students in the PTP condition seem to be disadvantaged in the simple task, compared to those in the OLP and CG conditions, as in the PTP they did not perform as well as those either in the CG in terms of WT/√2W and accuracy or in the OLP in terms of WT/√2W. PTP participants are supposed to write fast with limited attentional resources devoted to the *translating* process to retrieve advanced or various words during writing; in comparison, OLP and CG writers have enough time to search lexical resources. In other words, students under the PTP condition may experience a higher level of anxiety at the Output stage than those under the OLP and CG conditions, inhibiting pre-task planners to retrieve lexical items. This may explain the poorer performance of the PTP-H group in lexical diversity (i.e., WT/√2W) than that of the OLP-H and CG-H groups in the simple task.

Participants under the CG condition, as discussed in Section 6.3.1.3, may prioritize their attention to accuracy. This is especially true for the high-anxiety writers. With a low risk-taking ability, high-anxiety CG learners are more likely to have control over the existing interlanguage system instead of taking risk. For instance, high-anxiety CG interviewee #02 stated that he tried to reach “a 100% correctness”. In contrast, high-anxiety PTP participants may have limited attention allocated to monitor the accuracy during the 25-minute Output stage, as anxiety can reduce the attention available to the task (Derakshan & Eysenck, 2009; MacIntyre & Gardner, 1994; Trebits, 2016). Probably, for this reason, high-anxiety participants under the CG condition, in the simple task, significantly outperformed those under the PTP condition in terms of accuracy.

For low-anxiety participants, the disadvantage of the PTP group in lexical diversity and accuracy disappeared, with no significant differences found among the three planning groups in the simple task. A low level of anxiety may help pre-task planners to reduce their pressure on the transfer of the plan at the Output stage, making up for PTP students' shortcomings in accuracy and lexical performance. Alternatively, it may also be the case that low-anxiety students are less likely to be severely influenced by certain situations as their anxiety level is relatively stable. A low level of anxiety lowers the affective filter, ensuring their performance is representative of their actual writing competence. It is, therefore, not surprising that the low-anxiety participants, who shared a similar writing proficiency (see results of pre-test in Section 4.3), had a similar performance across the three planning groups in the simple task.

No significant differences were identified across the three planning conditions in the high-anxiety group with the complex task. As revealed in the interview, high anxiety-prone individuals' anxiety, regardless of planning conditions, is rapidly and severely raised by the cognitively taxing task. This may restrict high-anxiety writers benefitting from the cognitive-task variability manipulation as well as planning conditions, leading to similar performance.

For the complex task, the results showed that the PTP-L group outperformed the OLP-L group in content (d=1.055) and organisation (d=1.159), and CG-L participants wrote more accurately than OLP-L writers (d=.955 for Err/T, d=.915 for EP100). A similar pattern of results for the complex task was also found for RQ2. The OLP group was surpassed by the PTP group in content (d = .623) and organisation (d=.602), and by the CG in accuracy (d = .423 for Err/T, d = .477 for EP100). Larger effect sizes were elicited from the CG-L, OLP-L and PTP-L groups than from the CG, OLP and PTP groups in Stage I of Phase II, indicating that a low level of anxiety reinforced the effects of planning conditions on the complex task performance. This tendency is congruent with Robinson's (2011a) CH that learners' affective factors will more clearly differentiate learners' performance in the complex tasks.

The low-anxiety participants under planning conditions including OLP (in MLC and CP/C) and PTP (in MLC) appeared to have marked advantages with syntactic complexity over those under the CG condition when completing the complex task, which has not been found for RQ2. As discussed in Section 6.3.3.3, participants in the

CG seem to devote their attention primarily to accuracy at the expense of the linguistic complexity. The low-anxiety students with the help of planning (i.e., OLP and PTP) are likely to take risks, which may encourage them to prioritize attentional resources to the restructuring process and produce more complex syntactic frames (Skehan, 1996b). This may explain their better performance in syntactic complexity in the PTP-L and OLP-L groups than the CG-L group. It should be, however, noted that such a result was not observed in the interview for the low-anxiety participants under the PTP and OLP conditions, in which the pre-task and online planners did not explicitly refer to their confidence, or advantages in syntactic performance. Some potential factors may account for such gap: Participants' instant awareness of such advantage was not successfully captured by the post-task interview because of the time gap or, in the small sample for interview, there were no participants for whom this occurred; finally, Chinese students may be too modest to express their advantages during the interview. Further research is, therefore, warranted to elicit participants' instant feelings of anxiety and allocation of attention by using think-aloud methods for synchronous tracking or stimulated recall with video recording.

6.5 Proposal for a New Model: A Model of Cognitive Task Complexity in EFL Writing Processes

The varying pattern of results from the three planning groups in the present research and the inconsistent findings of the effects of task complexity in the L2 writing field in the prior studies (Cho, 2015; Frear & Bitchener, 2015; Johnson et al., 2012; Kormos, 2011; Kuiken & Vedder, 2007b, 2008, 2012; Ong & Zhang, 2010, 2013; Ruiz-Funes, 2015), suggest that L2 writing production may not be accurately described by LACM, nor CH. It seems that these two L2 oral models may not be suitable for the L2 writing field, because writing, as a recursive process, is fundamentally different from speaking, a linear process (Johnson et al., 2012). Therefore, a model of cognitive task complexity in relation to writing is proposed. This model, the Model of Cognitive Task Complexity in EFL Writing Processes, is based on the empirical results of this study as well as Kellogg's writing model (1996), Robinson's CH and Skehan's LACM.

The Model of Cognitive Task Complexity in EFL Writing Processes (see Figure 6.1) illustrates the effects of cognitive task complexity on EFL writing processes and performance with the inclusion of cognitive-task variability, planning type, and affective factors. In this model, the writing processes in Kellogg's writing model (1996) have been extended in a more detailed way, and relations between writing processes and performance have been built to predict the effects of cognitive task complexity on EFL writing. The present research focuses mainly on the cognitive writing processes, and the slave systems of visuospatial sketchpad and phonological loop in Kellogg's writing model (1996) that are not the researcher's concern are not discussed here.

The cognitive task complexity is hypothesized to be "inherent in the formulation stage", which is "determined by the demand tasks make on the planning of the content of the written text and/or on the linguistic encoding of the content" (Kormos, 2011, p. 150). As presented in Figure 6.1, the bolded "Formulation (CTC)" refers to that cognitive task complexity is inherent in the formulation stage. The *planning* sub-process is further divided into: *macro-structure planning* and *micro-structure planning*. *Macro-structure planning* mainly concerns the conceptualization of ideas during writing, which is associated with the pragmatic goals of the tasks, and closely related to content and organisation performance. The *macro-structure planning* is commonly accompanied by *micro-structure planning*, during which EFL writers may do some quick topic-related linguistic preparation, like lexis preparation. The linguistic preparation at this stage is fragmentized and tentative, which is prepared for but differs from the *translating* process. In *translating* process, writers systematically translate ideas by selecting lexical units and building syntactic frames into complete sentences. *Translating* is specialized as *lexical translating* and *syntactic translating*. The whole *translating* process and a small part of the *planning* process (i.e., micro-structure *planning*) are linked to restructuring and complexifying the interlanguage. Writers' lexical complexity mainly depends on what and how the lexical units are encoded during *lexical translating* process. *Syntactic translating* is responsible for syntactic complexity. *Monitoring* primarily concerns the control over of the existing interlanguage system, which is closely related to writers' accuracy. Apart from the local aspects, global aspects, such as paragraph and text organisation problems are also monitored by writers at this stage. Finally, fluency depends on the extent to which writers evolve in the three writing processes: formulation, execution and *monitoring*. The dotted line shown in Figure 6.1 illustrates the potential relations between writing processes and writing performance. The arrows present the writing information, as predicted in Kellogg's model (1996), flows from formulation to execution and forward to *monitoring*. The simultaneous activation of formulation, execution, and *monitoring* and their interaction also exist, as shown by the bidirectional arrows. For example, if writers struggled to translate ideas into sentences or lemmas, the *monitoring* process would reject the outputs and boost further *planning*.

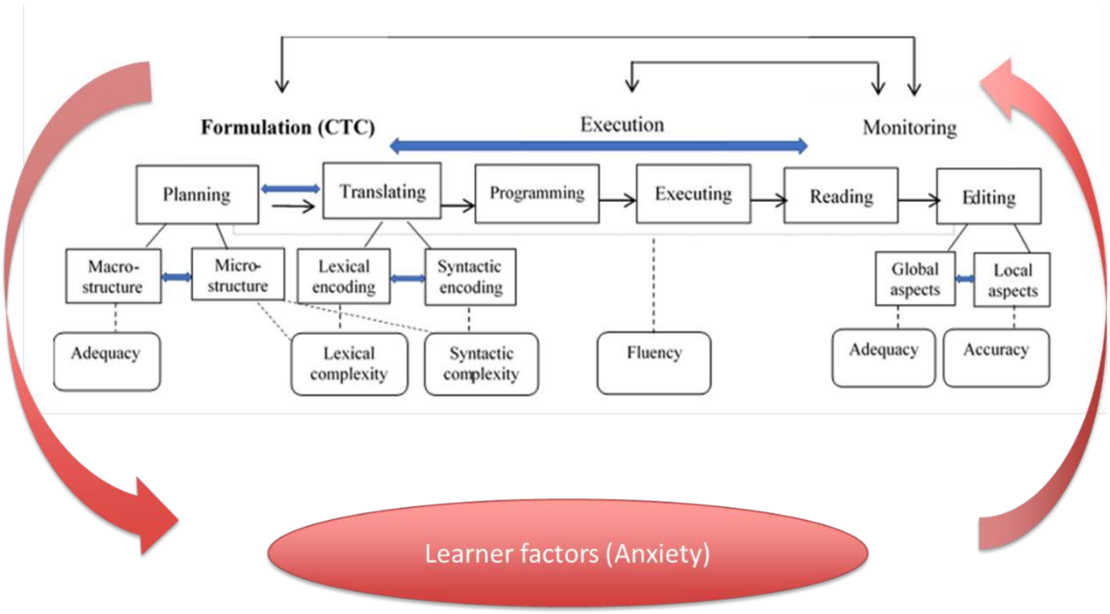

Figure 6.1 A Model of Cognitive Task Complexity in EFL Writing Processes (Based on Kellogg's writing model (1996))

Kellogg (1996) claimed that the central executive, involved in all the sub-processes with the exception of *executing*, is limited; Skehan (2013) similarly stated that humans have a limited information processing capacity. The blue bold bidirectional arrows refer to the potential trade-offs between the two related processes; that is, between formulation and *monitoring*, between *planning* and *translating*, between *lexical translating* and *syntactic translating*, between *macro-structure planning* and *micro-structure planning*, and between *global monitoring* and *local monitoring*. During the *planning* process, if EFL writers do both the macro-level and micro-level planning, tensions might be raised at the *planning* stage. During the *translating* process, both lexical and syntactic items are retrieved to encode ideas into text and there might be a trade-off between syntactic and lexical means of expression, when learners experience pressure at the *translating* stage. Within the formulation process, *planning* and *translating* may vie for the attentional resources when learners' attentional limits are reached for cognitively demanding task. For example, the linguistic complexity may be achieved at the cost of writers' tailoring their planned ideas and content; or pragmatic goals come first, with writers' failing to use their available advanced linguistic resources. This may explain why trade-offs have been found between the local linguistic complexity and the higher-order adequacy in some studies (Kuiken & Vedder, 2008, 2011; Pallotti, 2009). During the *monitoring* process, if writers allocate more attention to local accuracy, less attention would be directed to the global *monitoring*. Moreover, the trade-off between complexity and accuracy and between content and form predicted by Skehan (2013) may be due to the competition for attentional resources between the formulation and *monitoring* processes.

This model explains how different conditions impose a range of cognitive demands on the different writing processes, leading to varying writing performance. Looking back in the present study, the results for the PTP group indicate that if pre-task planning is provided, students may do more preparation for the complex task, than the simple task, during the *planning* stage and perhaps *translating* stage, ease the processing burden and enable more attention to be channelled into formulation and *monitoring* processes during writing. In the OLP group, increasing cognitive-task variability may trigger a tension between *planning* and *translating* process, and between *lexical translating* and *syntactic translating* sub-processes. This explains why a trade-off between adequacy and syntactic complexity, and between syntactic complexity and lexical complexity were identified for the OLP group in the complex task. In the comparison group, increased cognitive-task variability may lead to competition for attentional resources between formulation and *monitoring*, and between *local monitoring* and *global monitoring*, with the consequence an increase in accuracy, but a decrease in adequacy and lexical complexity. Compared to the CG condition, pre-task planning may facilitate the *planning* process and, accordingly, the performance in adequacy, while online planning may primarily assist the *translating* process and, consequently, the linguistic complexity performance. Learners seem to allocate more attentional resources to *monitoring* when writing in their own style (i.e., CG), resulting in better performance in accuracy. The OLP group might also allocate some working memory to the *monitoring* process, but pre-task planners are likely to have limited amount of time devoted to *monitoring*, particularly local aspects, compared to the OLP and the CG in the present study.

This model also includes learner factors. The orange bold bidirectional arrows shown in Figure 6.1 represent a non-linear relationship between cognitive task complexity and writing anxiety, one of the learner factors. Anxiety, cognitive-task variability and planning type, interact to influence students' L2 writing process and performance. Different levels of cognitive-task variability and varying planning types, to some extent, arouse or reduce participants' anxious feelings, consequently influencing students' attention allocation during writing process and, then, writing performance. Meanwhile, students' anxiety levels, more or less influence their perceptions of cognitive-task variability and different planning conditions, which further affect writers' attention allocation at each writing process and, finally, their writing performance.

The Model of Cognitive Task Complexity in EFL Writing Processes fills a research gap by combining cognitive task complexity, writing processes and writing performance as a whole. As discussed in Chapter 2, the existing models have not addressed the issue of how cognitive load, imposed by task complexity, affects EFL writing process and performance. For instance, Kellogg's writing model does not consider the role of task environment and affective factors in the writing process, nor predict performance outcomes of manipulating task design and implementation factors. The two influential models, LACM and CH, predict the effects of task complexity on oral production, which may not be exactly suitable for the L2 writing field (Johnson et al., 2012). In addition, the inclusion of cognitive-task variability, planning type, and affective factors makes the proposed model closely approximate the real writing situation. For example, pre-task and online planning are strategies writers often use during writing, and writing anxiety is what L2 learners frequently experience in the composing process.

6.6 Chapter Summary

This chapter interpreted the findings concerning relevant empirical studies and theoretical frameworks.

Increases in cognitive-task variability resulted in enhancements in accuracy, fluency adequacy and perhaps, syntactic complexity under the PTP condition, which is consistent with Robinson's CH. Increasing cognitive-task variability led to trade-offs between syntactic and lexical complexity, and between syntactic complexity and adequacy under the OLP condition, and trade-offs between lexical complexity and accuracy, and between accuracy and adequacy under CG, which supports Skehan's LACM. Furthermore, the results suggest that pre-task planning mainly assists the *planning* stage, online planning primarily facilitates the *translating* process, and students writing in their own way are likely to focus on *monitoring* process. Writing anxiety is quantitatively and qualitatively found to interfere with the effects of cognitive task complexity on L2 writing performance. Based on these results, the Model of Cognitive Task Complexity in EFL Writing Processes is proposed to illustrate the effects of anxiety, cognitive-task variability and planning type on EFL writing process and performance.

In the next chapter, conclusions, implications and limitations of the study will be stated, and the directions for further research will be suggested.

Chapter 7 Conclusion

7.1 Overview

In this chapter, the main findings related to the research questions are summarized, followed by an elaboration of the theoretical, methodological and practical implications of this study. The limitations and recommendations for future research conclude the chapter.

7.2 Summary of Major Findings

The three main objectives of the study were to: (a) validate the cognitive-task variability manipulated within the two argumentative tasks, (b) explore the isolated and synergistic effects of cognitive-task variability and planning type on Chinese EFL students' argumentative writing, and (c) examine how writing anxiety interacts with cognitive task complexity to influence Chinese EFL students' writing. Major findings in relation to the three sets of research questions are presented.

7.2.1 Validation of cognitive-task variability

Students rated the designed-to-be complex task more difficult and consuming greater mental-effort than the simple version, based on the TDMEQ. In dual-task Experiment 1, participants achieved significantly more accurate rates on the secondary task when carrying out the simple primary writing task than the complex version. In dual-task Experiment 2, participants' secondary-task reaction time for the complex primary writing task was significantly longer than for the simple primary writing task.

7.2.2 Effects of cognitive-task variability and planning type

The interaction was significant between cognitive-task variability and planning type in terms of lexical complexity (i.e., $WT/\sqrt{2W}$), fluency, and adequacy (i.e., content, organisation and overall writing quality).

As for the PTP group, increases in cognitive-task variability did not influence participants' lexical complexity, but positively affected syntactic complexity in terms of CP/C, accuracy in terms of EP100, fluency, and adequacy in terms of content, organisation and overall writing quality. With the OLP group, increasing cognitive-task variability led to no effects on accuracy and fluency, but decreases in lexical complexity (i.e., WT/√2W) and adequacy (i.e., content, organisation and overall writing quality). A two-way influence on syntactic complexity was identified for the OLP group with reductions in C/T and DC/C and rises in CP/C, as a function of increasing cognitive-task variability. For the CG, better performance in lexical complexity (i.e., LD and WT/√2W) and adequacy (i.e., content, organisation and overall writing quality) was found in the simple task than the complex task version, with greater accuracy (i.e. Err/T) identified in the complex task than in the simple one.

In the simple task, the PTP group was outperformed by the OLP group in syntactic complexity (i.e., DC/C) and lexical complexity (i.e., WT/√2W), and by the CG in lexical complexity (i.e., WT/√2W) and accuracy (i.e., Err/T, EP100 and EFC/C). In the complex task, the PTP group surpassed the CG in fluency and adequacy (i.e., content, organisation and writing quality) and the OLP group in content and organisation. The OLP group had the advantages over the CG in lexical complexity (i.e., WT/√2W), and fluency. The CG gained a better position in accuracy (i.e., Err/T and EP100), compared to the PTP and OLP groups.

7.2.3 Relationships among anxiety, cognitive-task variability, planning type and EFL writing performance

The results obtained from the quantitative (Stage II of Phase II) and the qualitative data were consistent. As identified in the interviews, participants from the high-anxiety group seem to be more prone to L2 writing anxiety than those from the low-anxiety group.

The quantitative analysis found no significant interaction of task*group*anxiety on writing performance, but anxiety interfered with cognitive task complexity to influence students' writing performance:

(a) For the high-anxiety group, increasing cognitive-task variability led to significant decreases in syntactic complexity and lexical complexity under the OLP condition, syntactic complexity under PTP, and lexical complexity and adequacy under CG, but increases in accuracy and fluency under PTP. In contrast, increases in cognitive-task variability substantially improved low-anxiety students' EFL writing in syntactic complexity under the OLP condition, adequacy and fluency under PTP, and accuracy performance under CG.

(b) Low-anxiety PTP students significantly outperformed their high-anxiety counterparts in lexical complexity and overall writing quality in the simple task, and in syntactic complexity, lexical complexity and adequacy in the complex task. These patterns were, however, not identified in either the OLP or the CG conditions in both simple and complex task versions.

(c) In the simple task, the PTP-H group was significantly outperformed by the CG-H group in lexical complexity and accuracy and by the OLP-H group in terms of lexical complexity. In the complex task, syntactic complexity of the OLP-L and of the PTP-L significantly outnumbered that of the CG-L. Participants in the PTP-L group had greater advantages in adequacy performance and the CG-L in accuracy over the OLP-L group.

The qualitative data revealed a non-linear relationship among anxiety, cognitive-task variability, and planning type, which confirms the quantitative findings that these three variables interacted with each other to influence students' EFL writing performance. On the one hand, different levels of cognitive-task variability and varying planning types arouse or reduce participants' anxious feelings, to some extent, and consequently influence students' writing process and writing performance. On the other hand, students' anxiety levels influence their perceptions of cognitive-task variability and different planning conditions, and then writing process and performance.

7.3 Theoretical Implications

This research contributes empirical evidence to the effects of cognitive task complexity (along each and both dimension(s)) on students' EFL writing. Increasing cognitive-task variability under the OLP condition led to trade-offs between syntactic and lexical complexity, and between syntactic complexity and adequacy. Trade-off effects were also identified for the CG between lexical complexity and accuracy, and between accuracy and adequacy, when the cognitive-task variability was increased. The results for the OLP and CG conditions indicate that competition for learners' allocation of attention may exist between the local linguistic perspective and the higher-order process, and within the local aspects. These findings support the basic principle of Skehan's (2013) LACM that "humans have a limited information processing capacity and must therefore prioritise where they allocate their attention" (p. 189). With the PTP condition, increasing cognitive-task variability and providing pre-task planning enhanced accuracy, fluency adequacy, and perhaps, syntactic complexity. These findings are congruent with one of Robinson's predictions that simultaneously increasing task complexity along resource-directing dimensions and reducing the complexity of resource-dispersing task characteristics will affect learners' EFL performance positively.

Furthermore, the results reveal that different planning types ease the pressure on the central executive in working memory in different ways, leading to different effects on learners' allocation of attention and EFL production. For instance, OLP and PTP facilitate the formulation process but in different ways, compared to CG. Pre-task planning appears to facilitate the *planning* process, while online planning may assist primarily the *translating* process. These results not only confirm Ellis and Yuan's (2004) prediction that different planning types have different effects on students' writing quality, but also suggest that Kellogg's writing model (1996) offers a suitable starting point for explaining the complex nature of writing processes in the task-based L2 writing research.

The following findings related to anxiety are consistent with Robinson's (2011a) CH that learners' affective factors (i.e., writing anxiety) contribute to perceptions of task difficulty, which will more clearly differentiate learners' performance in the complex tasks.

(a) Contrasting patterns of results for the high- and low-anxiety groups in relation to the effects of cognitive-task variability on writing performance were identified. For the high-anxiety group, increasing cognitive-task variability generally led to significant decreases in L2 writing under the OLP, PTP and CG conditions. By contrast, increases in cognitive-task variability substantially improved low-anxiety students' writing performance under the three planning conditions.
(b) Low-anxiety pre-task planners performed significantly better than their high-anxiety counterparts in writing performance, and more obviously in the complex task than the simple version.
(c) In the complex task, results for group comparison in accuracy and adequacy showed that larger effect sizes were elicited from the CG-L, OLP-L and PTP-L groups than from the CG, OLP and PTP groups in Stage I of Phase II, which suggests that a low level of anxiety reinforced the effects of planning conditions on the complex task performance.

This study, based on the findings, proposed the Model of Cognitive Task Complexity in EFL Writing Processes (see Section 6.5). This model extended Kellogg's writing model to address the main concern of Skehan's LACM and Robinson's CH in the EFL writing process, that is, how learners' attentional resources are actually allocated during task completion. This model dynamically predicts the effects of anxiety, cognitive-task variability and planning type on EFL writing processes and performance. Learners have limited attentional resources and the competition for attentional resources may exist between formulation and *monitoring*, between *planning* and *translating*, between *lexical* and *syntactic translating* processes, between *macro-structure planning* and *micro-structure planning*, and between *global monitoring* and *local monitoring*. With different task-completion conditions provided, learners might differentially prioritise their allocation of attention, leading to different writing performance. Anxiety may interact with cognitive task complexity to influence students' writing process and performance. For example, the provision of online planning may direct attention to *translating*, enhancing writing performance in terms of syntactic and lexical complexity. When cognitive-task variability was increased for learners under the OLP condition, tensions between *planning* and *translating*, and between lexical and syntactic *encoding* were triggered. This may explain why trade-offs were found between syntactic complexity and adequacy, and between lexical and syntactic complexity, in the present study, for the OLP group, when they completed the cognitively demanding task. This model fills gaps in LACM and CH, two influential oral models, which may not be exactly suitable to the L2 writing field; and in Kellogg's writing model in which it does not consider the role of cognitive task complexity nor the connections between writing process and performance.

7.4 Methodological Contributions

A methodological contribution of this study is that the findings highlight the necessity of validating the manipulation of cognitive-task variability in the writing field by using multiple measures of cognitive demands. The current study is the first that attempts to measure the cognitive load imposed by the writing tasks in the field of task-based research by using dual-task technique, although some researchers (e.g. Révész et al., 2016; Sasayama, 2016) have pioneered such technique in the oral tasks. Dual-task techniques, as well as self-ratings, are proven as reliable methods for validation of the cognitive-task variability in writing. Various cognitive demand measures have their own advantages and disadvantages and may capture distinct aspects of the cognitive load. Triangulation of these data sources is critical to ensure that the findings for the manipulations of cognitive-task variability, a key factor to cognitive task complexity, are accurate and reliable.

Another methodological contribution is the role of multi-dimensional linguistic measures and adequacy used in the present study. For example, choosing multi-dimensional syntactic complexity measures, rather than redundant indexes of a certain sub-construct, is essential to capture syntactic development of learners at different proficiency levels. This also applies to other linguistic measures, such as lexical complexity and accuracy. The inclusion of adequacy is also necessary to examine to what extent students have achieved the pragmatic goals of the writing tasks.

Finally, a mixed-methods design is crucial to provide an in-depth understanding of the role of writing anxiety in cognitive task complexity. The researcher made an effort to use a qualitative approach to provide detailed information on students' feelings and perceptions under different task conditions as well as to explain and complement the quantitative data. For instance, the interview in this study has captured information that the statistics and numbers cannot, such as, the sources of participants' L2 writing anxiety, perceptions of language anxiety related to cognitive-task variability and planning type, and learners' focus during planning and writing under different conditions. These qualitative data confirm, and illustrate, the quantitative findings that anxiety interfered with cognitive task complexity to influence writing performance.

7.5 Practical Implications

The results of the current study have some implications for EFL writing teachers as well as for curriculum designers in Chinese universities, including how to validate designed writing tasks, grade and sequence tasks by varying cognitive-task variability and planning type, and pay attention to the role of anxiety in EFL writing teaching.

The findings of this study underline the importance of validating designed writing tasks. EFL writing teachers and curriculum designers used to assume that the designed tasks could yield distinguished levels of cognitive demands. The *Task Difficulty and Mental Effort Questionnaire* (TDMEQ) has been shown in this study to be a valid and reliable tool for the assessment of cognitive demands. It is easy for teachers to use the TDMEQ to understand what learners' perceptions of task difficulty and mental effort are and whether the complexity of a task is suitable to learners. Accordingly, teachers can adjust the complexity and difficulty of writing tasks and their teaching plans. EFL learners' perceptions and feedback gathered by the TDMEQ can also guide the curriculum designers to design writing tasks with varying levels of cognitive demands and then evaluate the actual cognitive demands imposed by the designed tasks.

The findings for the effects of cognitive-task variability and planning type can provide guidance on making task-grading and sequencing decisions for EFL teachers when they schedule a syllabus for a given course. For example, increasing the cognitive-task variability leads to greater effort to conceptualize the ideas at the *planning* process. Consequently, complex linguistic structures would be used to encode complex ideas, increasing the demands in the *translating* process. Different planning types may ease learners' pressure on the formulation process in different ways, leading to different effects on learners' allocation of attention and EFL production; pre-task planning appears to facilitate the *planning* process, while online planning may assist the *translating* process primarily. Teachers can increase the cognitive-task variability of the writing tasks by involving more elements and reasoning demands. At the same time, teachers can teach EFL learners to use pre-task planning and online planning to efficiently identify ideas, generate and elaborate ideas, prime lexical elements and prepare syntactic frames, and carefully select different planning types for distinctive situations. In this way, learners could be encouraged to use more complex linguistic structures and improve their writing adequacy.

The teaching of writing is a challenging task for many EFL teachers, as scholars have documented (e.g., Wang & Zhang, 2019; Zhang, 2016; Zhang, Wang, & Wu, 2017). Results in this study may help teachers understand the allocation of learners' attentional resources during writing, and how different task features and implementation conditions may vary learners' attention available. Based on this knowledge, teachers can introduce the cognitive processes involved in writing to learners in order to raise their awareness of how these writing processes work together and guild learners to consciously allocate their attention during writing. Of note, some teaching activities, informed by the findings, may potentially influence students' attention allocation during writing. For example, teachers' feedback preference for accuracy may encourage learners to pay more attention to controlling their existing interlanguage system rather than complexifying linguistic encoding. Teachers should provide multi-dimensional written feedback in terms of complexity, accuracy, fluency, and adequacy to guide learners to allocate their attention to all the dimensions of writing performance. This is also true in the field of EFL writing assessment. Raters used to separately use CAF or (holistic) writing scores to evaluate students' writing performance during the writing tests. The current study highlights the importance of using a multi-dimensional criterion, including linguistic measures and adequacy performance, to assess learners' writing performance.

Furthermore, this study found anxiety interacts with cognitive task complexity to influence learners' writing performance. Anxiety, however, has been commonly overlooked during daily teaching. The results in the current study indicated that high-anxiety participants benefited less from the increased cognitive-task variability than their low-anxiety counterparts. Teachers can, therefore, use *the Second Language Writing Anxiety Inventory* (SLWAI) and post-task interviews to understand EFL learners' anxious feelings on the writing tasks and offer extra support for high-anxiety learners.

7.6 Limitations and Recommendations

7.6.1 Limitations

This study validated the manipulation of cognitive-task variability based on the data from Chinese EFL learners. No baseline data were, however, collected from native speakers of English, who have few language barriers when writing compositions in English. Future studies can extend Phase I of the present study by tracking native speakers' response to the tasks and use the two sources of the data (from L1 speaker and L2 speaker) to triangulate the manipulation of the cognitive-task variability.

Another limitation is that this study adopted convenience sampling rather than random sampling. The generalisability of the present findings is restricted to English majors in a Chinese-university context, despite attempts to enhance the external validity through collecting data from a large sample size. The issue of assigning intact classes as a whole into different experimental groups posed a potential threat to the external validity of the results, although participants' writing proficiency of each group was strictly controlled.

In future studies, the participant pool should be expanded to include a wider range of ages, majors and EFL ethnicities.

Information on what writers with different anxiety levels actually do and feel under different conditions was gathered through interview. The interviews, conducted after the completion of the second writing task, cannot capture the instant thought processes; hence, data from completing the first task might be lost due to recall failure. Think-alouds can be involved in future studies to tap students' moment-to-moment feelings and actions at different planning and writing processes (Zhang & Zhang, 2020).

7.6.2 Recommendations for future research

The current study is the first that attempted to validate the manipulations of cognitive-task variability in the writing field by using multiple measures of cognitive demands. Replication of this research using self-ratings and, in particular, dual-task methods is worthwhile. Other techniques, such as time estimation (Sasayama, 2016), eye-movement recordings (Révész, 2014), heart rate test and neuroimaging techniques, are also warranted to be involved in future studies to assess the cognitive demands in EFL writing.

This study investigated [±few elements] and [±reasoning demands] as resource-directing variables and pre-task planning and online planning as resource-dispersing variables. This research could be replicated in the future research, as there are a limited number of studies focused on the effects of cognitive task complexity simultaneously along the two dimensions on EFL writing. This study could also be extended by further cognitive task complexity related research, involving other variables, such as different proficiency levels, task types and task modality. Apart from the variables explored in the present study, many other variables predicted within Robinson's CH, such as, [±Here-and-Now] and [± prior knowledge], and Skehan's LACM, such as familiarity of the topic, are under-researched. Further research should investigate the effects of cognitive task complexity manipulated by other variables to determine the role of other aspects of cognitive task complexity in EFL writing.

This study tested the role of writing anxiety in cognitive task complexity. Additional studies are needed to detect how other affective factors, such as motivation, self-efficacy, working memory and aptitude, affect cognitive task complexity and EFL writing. To explore the effects of individual differences and cognitive task complexity on attention allocation during writing, follow-up open-ended questionnaires can be assigned to all participants to gain large sample-sized data for tracking what writers actually do, think and feel during writing. Think-aloud methods can also be used to tap students' real-time thought processes at different stages of writing.

7.7 Conclusion

This study was designed to investigate the effects of cognitive-task variability, planning type and writing anxiety on EFL writing drawing on Robinson's CH, Skehan's LACM and Kellogg's Writing Model. Results showed that the manipulations of cognitive-task variability in this study proved to be successful with the designed-to-be complex task version being more cognitively demanding and difficult. Furthermore, increases in cognitive-task variability resulted in enhancements in accuracy, fluency adequacy and perhaps, syntactic complexity under the PTP condition. Trade-off was identified, as a function of increasing cognitive-task variability, between the local linguistic perspectives as well as between the local and the higher-order process under both the OLP and CG conditions. Moreover, different planning types were found to help ease the pressure on working memory in different ways, leading to varying effects on learners' attention allocation. Finally, writing anxiety was found, quantitatively and qualitatively, to interfere with the effects of cognitive task complexity on EFL writing performance. Based on the empirical findings and existing models (i.e., CH, LACM, and Kellogg's writing model), the current study proposed the Model of Cognitive Task Complexity in EFL Writing Processes. In addition to the contribution to the knowledge of the effects of cognitive task complexity on EFL writing, this study has implications for the future research and teaching activities related to task-based syllabus construction, writing assessment as well as understandings of the nature of EFL learners' attentional resources.

References

Abrams, Z. I. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task-complexity. *System*, *81*, 110–121. https://doi.org/10.1016/j.system.2019.01.009

Abrams, Z. I., & Byrd, D. R. (2016). The effects of pre-task planning on L2 writing: Mind-mapping and chronological sequencing in a 1st-year German class. *System*, *63*, 1–12. https://doi.org/10.1016/j.system.2016.08.011

Andrews, R. (1995). *Teaching and learning argument*. London: Cassel.

Arnold, J. (2000). *Affect in language learning*. Beijing: Foreign Language Teaching and Research Press.

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baralt, M., Gilabert, R., & Robinson, P. (2014a). An introduction to theory and research in task sequencing and instructed second language learning. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 1–34). London, UK: Bloomsbury.

Baralt, M., Gilabert, R., & Robinson, P. (2014b). *Task sequencing and instructed second language learning*. London, UK: Bloomsbury.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5–35. https://doi.org/10.5054/tq.2011.244483

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 53–61. https://doi.org/10.1207/S15326985EP3801

Brünken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 181–202). Cambridge, UK: Cambridge University Press.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, *49*(2), 109–119. https://doi.org/10.1027//1618-3169.49.2.109

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, *26*, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Bygate, M. (Ed.). (2015). *Domains and directions in the development of TBLT: A decade of plenaries from the international conference*. Amsterdam, Netherlands: Benjamins.

Calbraith, D., & Torrance, M. (2004). Revision in the context of different drafting strategies. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 63–85). Boston: Kluwer.

Carroll, J. B. (1967). On sampling from a lognormal model of word frequency distribution. In H. Kučera & W. N. Francis (Eds.), *Computational analysis of present-day American English* (pp. 406–424). Providence, RI: Brown University Press.

Carter, R. (1998). *Mapping the mind*. London, UK: Weidenfeld & Nicolson.

Cheng, Y. S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing*, *13*(4), 313–335. https://doi.org/10.1016/j.jslw.2004.07.001

Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, *49*(3), 417–446. https://doi.org/10.1111/0023-8333.00095

Cho, H. (2015). Effects of task complexity on English argumentative writing. *English Teaching*, *70*(2), 107–131. https://doi.org/10.15858/engtea.70.2.201506.107

Choi, J. (2013). Does peer feedback affect L2 writers' L2 learning, composition skills, metacognitive knowledge, and L2 writing anxiety? *English Teaching*, *68*(3), 187–

213. https://doi.org/10.15858/engtea.68.3.201309.187

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Creswell, J. W. (2013). *Research design : Qualitative, quantitative, and mixed methods approaches*. London, UK: Sage.

Cui, B. (2011). Validating a new measure of EFL learners' lexical sophistication. *Shandong Foreign Language Teaching Journal*, *6*, 44–52.

Daly, J. A., & Miller, M. D. (1975a). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, *9*(3), 242–249.

Daly, J. A., & Miller, M. D. (1975b). Further studies on writing apprehension: SAT scores, success expectations, willingness to take advanced courses and sex differences. *Research in the Teaching of English*, *9*(3), 250–256.

Derakshan, N., & Eysenck, M. W. (2009). Anxiety, processing efficiency, and cognitive performance: New developments from attentional control theory. *European Psychologist*, *14*(2), 168–176. https://doi.org/10.1027/1016-9040.14.2.168

Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Erlbaum.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford, UK: Oxford University Press.

Dörnyei, Z. (2010). *Questionnaires in second language research : Construction, administration, and processing*. New York, NY: Routledge.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.

Ellis, R. (2005). *Planning and task performance in a second language*. Amsterdam, Netherlands: Benjamins.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, *30*(4), 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R., Skehan, P., Li, S. F., Shintani, N., & Lambert, C. (2019). *Task-based language teaching : Theory and practice*. Cambridge, UK: Cambridge University Press.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and acuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*(1), 59–84. https://doi.org/10.1017/S0272263104261034

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis (5th ed)*. Hoboken, NJ: Wiley.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Field, A. P. (2009). *Discovering statistcs using SPSS: (and sex, drugs and rock "n" roll)* (3rd ed.). London, UK: Sage.

Flower, L., & Hayes, J, R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication*, *1*(1), 120-160. https://doi.org/10.1177/0741088384001001006

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, *3*(3), 215–247. https://doi.org/10.1177/136216889900300303

Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: The effects on accuracy, complexity, and fluency of second language performance. *Canadian Modern Language Review*, *69*(3), 249–273. https://doi.org/10.3138/cmlr.69.3.249

Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, *30*, 45–57. https://doi.org/10.1016/j.jslw.2015.08.009

Freedman, A., & Pringle, I. (1984). Why students can't write arguments. *English in Education*, *18*(2), 73–84. https://doi.org/10.1111/j.1754-8845.1984.tb00668.x

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, UK: Erlbaum.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and [±Here-and-Now]: Effects on L2 oral production. In M. D. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44–68). Clevedon, UK: Multilingual Matters.

Guo, Y., & Qin, X. (2010). An empirical study on foreign language writing anxiety of Chinese non-English majors. *Foreign Language World*, *2*, 54–82.

Halliday, M. A., & Mattiesen, C. (1999). *Construing experience through meaning: A language-based approach to cognition*. London, UK: Continuum.

Hirose, K. (2003). Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing*, *12*(2), 181–209. https://doi.org/10.1016/S1060-3743(03)00015-8

Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*(1), 112–126. https://doi.org/10.1017/S0267190501000071

Horwitz, E. K. (2016). Factor structure of the foreign language classroom anxiety scale: Comment on Park (2014). *Psychological Reports*, *119*(1), 71-76. https://doi.org/10.1177/0033294116653368

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70(2)*, 125–132.

Horwitz, E. K., & Young, D. (1991). *Language anxiety*. Upper Saddle River, NJ: Prentice Hall.

Housen, A., De Clercq, B., Kuiken, F., & Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second Language Research*, *35*(1), 3–21. https://doi.org/10.1177/0267658318809765

Huang, Y., & Zhang, L. J. (2020). Does a process-genre approach help improve students' argumentative writing in English as a foreign language? Findings from an intervention study. *Reading & Writing Quarterly*, *36*(4), 339–364. https://doi.org/10.1080/10573569.2019.1649223

Hunt, K. (1966). Recent measures in syntactic development. *Elementary English*, *43*(7), 732–739.

Ishikawa, T. (2007). The effect of manipulating task complexity along the [±Here-and-Now] dimension on L2 written narrative discourse. In M. D. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Clevedon, UK: Multilingual Matters.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V. F., & Hughey, J. (1981). *Testing EFL composition: A practical approach*. Rowley, MA: Newbury.

Jin, L., & Cortazzi, M. (2002). English language teaching in China: A bridge to the future. *Asia Pacific Journal of Education*, *22*(2), 53–64. https://doi.org/10.1080/0218879020220206

Jin, Y., & Fan, J. (2011). Test for English Majors (TEM) in China. *Language Testing*, *28*(4), 589–596. https://doi.org/10.1177/0265532211414852

Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, *21*(3), 264–282. https://doi.org/10.1016/j.jslw.2012.05.011

Kang, S., & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing*, *45*, 61–72. https://doi.org/10.1016/j.jslw.2019.08.001

Kellogg, R. T. (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes. *Memory & Cognition*, *15*(3), 256–266.

Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *The American Journal of Psychology*, *103*, 327–342.

Kellogg, R. T. (1996). A model of writing memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Mahwah, NJ: Erlbaum.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research*, *5*(2), 159–190. https://doi.org/10.17239/jowr-2013.05.02.1

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, *20*(2), 148–161. https://doi.org/10.1016/j.jslw.2011.02.001

Krashen, S. (1982). *Principles and practive in second language acquisition*. Oxford, UK: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London, UK: Longman.

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. *EUROSLA Yearbook*, *5*(1), 195–222. https://doi.org/10.1075/eurosla.5.10kui

Kuiken, F., & Vedder, I. (2007a). Cognitive task complexity and linguistic performance in French L2 writing. In M. D. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 117–135). Clevedon, UK: Multilingual Matters.

Kuiken, F., & Vedder, I. (2007b). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, *45*(3), 261–284. https://doi.org/10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1), 48–60. https://doi.org/10.1016/j.jslw.2007.08.003

Kuiken, F., & Vedder, I. (2011). Task complexity and linguisitc performance in L2 writing and speaking. In P. Robinson (Ed.), *Second language task complexity : Researching the cognition hypothesis of language learning and performance* (pp. 91–104). Amsterdam, Netherlands: Benjamins.

Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 143–170). Amsterdam, Netherlands: Benjamins.

Kuiken, F., & Vedder, I. (2019). Investigating teachers' perceptions of syntactic complexity in L2 academic writing. *Instructed Second Language Acquisition, 3*(2), 228–248. https://doi.org/10.1558/isla.37977

Kuiken, F., Vedder, I., Housen, A., & De Clercq, B. (2019). Variation in syntactic complexity: Introduction. *International Journal of Applied Linguistics, 29*(2), 161–170. https://doi.org/10.1111/ijal.12255

Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, *17*(2), 69–85. https://doi.org/10.1016/j.jslw.2007.10.001

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, *63*(1), 13–22. https://doi.org/10.1093/elt/ccn010

Lee, I. (2011). Working smarter, not working harder: Revisiting teacher feedback in the L2 writing classroom. *Canadian Modern Language Review*, *67*(3), 377–399. https://doi.org/10.3138/cmlr.67.3.377

Liu, B., Wang, Y., & Zhang, L. J. (2017). Effects of task type on planning and writing production in online EFL writing. *Modern Foreign Languages*, *40*(1), 102–113.

Long, M. H. (1996). The linguistic environment for second language aquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, *26*(1), 27–56. https://doi.org/10.2307/3587368

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, *45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *Modern Language Journal*, *96*(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x

MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 24–45). Boston, MA: McGraw Hill.

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11-30). Bristol, UK: Multilingual Matters.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, *44*(2), 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

Madrid, R. I., Van Oostendorp, H., & Melguizo, M. C. P. (2009). The effects of the number of links and navigation support on cognitive load and learning with hypertext: The mediating role of reading order. *Computers in Human Behavior*, *25*(1), 66–75. https://doi.org/10.1016/j.chb.2008.06.005

Manchón, R. M., & De Larios, J. R. (2007). On the temporal nature of planning in L1 and L2 composing. *Language Learning*, *57*(4), 549–593. https://doi.org/10.1111/j.1467-9922.2007.00428.x

Mao, S., & Crosthwaite, P. (2019). Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practices. *Journal of Second Language Writing*, *45*, 46–60. https://doi.org/10.1016/j.jslw.2019.05.004

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*(1), 49–63. https://doi.org/10.1037//0022-0663.88.1.49

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3-15 https://doi.org/10.1016/j.jslw.2015.06.004

McKinley, J., & Rose, H. (2018). Conceptualizations of language errors, standards, norms and nativeness in English for research publication purposes: An analysis of journal submission guidelines. *Journal of Second Language Writing*, *42*, 1–11. https://doi.org/10.1016/j.jslw.2018.07.003

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes*, *47*(4), 292–330. https://doi.org/10.1080/01638530902959943

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83–108.

Michel, M., Révész, A., Shi, D., & Li, Y. (2019). The effects of task demands on linguistic complexity and accuracy across task types and L1/L2 speakers. In Z. E. Wen & M. J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy:*

In honour of Peter Skehan (pp. 133–151). Amsterdam, Netherlands: Benjamins. https://doi.org/10.1075/tblt.13.07mic

Norris, J. M., & Ortega, L. (2003). Defining and measuring SLA. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 717–761).Malden, MA: Blackwell.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System*, *34* (4), 566–585. https://doi.org/10.1016/j.system.2006.08.003

Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, *19*(4), 218–233. https://doi.org/10.1016/j.jslw.2010.10.003

Ong, J., & Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Quarterly*, *47*(2), 375–398. https://doi.org/10.1002/tesq.55

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, *21*(1), 109–148.https://doi.org/10.1017/S0272263199001047

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency : A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Oxford, R. L. (2000). Anxiety and the language learner: New insights. In J. Arnold (Ed.), *Affect in language learning* (pp. 58–67). Beijing: Foreign Language Teaching and Research Press.Paas, F. G. W. C, Van Merriënboer, J. J. G, & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419-430. https://doi.org/10.2466/pms.1994.79.1.419

Paas, F. G. W. C., Van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419-430. https://doi.org/10.2466/pms.1994.79.1.419

Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th edition). Sydney: Allen & Unwin..

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30*(4), 590–601. https://doi.org/10.1093/applin/amp045

Polio, C., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, *26*, 10–27. https://doi.org/10.1016/j.jslw.2014.09.003

Rahimi, M. (2016). *Task complexity, affective factors, and pre-task planning: Effects on L2 writing production*. Unpublished doctoral thesis, University of Auckland, Auckland, New Zealand. https://doi.org/10.1177/1097184X15574339

Rahimi, M. (2019). Effects of increasing the degree of reasoning and the number of elements on L2 argumentative writing. *Language Teaching Research*, *23*(5), 633–654. https://doi.org/10.1177/1362168818761465

Rahimi, M., & Zhang, L. J. (2018). Effects of task complexity and planning conditions on L2 argumentative writing production. *Discourse Processes*, *55*(8), 726-742. https://doi.org/10.1080/0163853X.2017.1336042

Rahimi, M., & Zhang, L. J. (2019). Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Reading and Writing*, *32*(3), 761–786. https://doi.org/10.1080/0163853X.2017.1336042

Read, J. (2000). *Assessing vocabulary*. Cambridge, UK: Cambridge University Press.

Reed, W. M., Burton, J. K., & Kelly, P. P. (1985). The effects of writing ability and mode of discourse on cognitive capacity engagement. *Research in the Teaching of English*, *19*(3), 283–297. https://doi.org/10.2307/40171051

Révész, A. (2011). Task Complexity, focus on L2 constructions, and individual differences: A classroom-based study. *Modern Language Journal*, 95, 162-181.https://doi.org/10.1111/j.1540-4781.2011.01241.x

Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics*, *35*(1), 87–92. https://doi.org/10.1093/applin/amt039

Révész, A., Kourtali, N. E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241. https://doi.org/10.1111/lang.12205

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, *38*(4), 703–737. https://doi.org/10.1017/S0272263115000339

Révész, A., Sachs, R., & Hama, M. (2014). The effects of task complexity and input frequency on the acquisition of the past counterfactual construction through recasts. *Language Learning*, *64*(3), 615–650. https://doi.org/10.1111/lang.12061

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge, UK: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43*(1), 1–32. https://doi.org/10.1515/iral.2005.43.1.1

Robinson, P. (2007a). Criteria for classifying and sequencing pedagogic tasks. In M. D.

P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 7–26). Clevedon, UK: Multilingual Matters.

Robinson, P. (2007b). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching, 45*(3), 193–213. https://doi.org/10.1515/iral.2007.009

Robinson, P. (2010). Situating and distributing cognition across task demand: The SSARC model of pedagogic task sequencing. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition : Inside the learner's mind* (pp. 243–269). Amsterdam, Netherlands: Benjamins.

Robinson, P. (2011a). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity : Researching the cognition hypothesis of language learning and performance* (pp. 3–38). Amsterdam, Netherlands: Benjamins.

Robinson, P. (2011b). Task based language learning: A review of issues. *Language Learning*, *61(s1)*, 1–36. https://doi.org/10.1111/j.1467-9922.2011.00641.x

Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 87–122). Amsterdam, Netherlands: Benjamins.

Robinson, P., & Gilabert, R. (2019). Task-based learning: Cognitive underpinnings. In C. A. Chapelle (Ed.), *The concise encyclopedia of applied linguistics* (pp. 1–6). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781405198431.wbeal1143.pub2

Ruiz-Funes, M. (2014). Task complexity and linguistic performance in advanced college-level foreign language writing. In H. Byrnes & R. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 163–192). Amsterdam, Netherlands: Benjamins.

Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing*, *28*, 1–19. https://doi.org/10.1016/j.jslw.2015.02.001

Sasayama, S. (2016). Is a ‘complex’ task really complex? Validating the assumption of cognitive task complexity. *Modern Language Journal*, *100*(1), 231–254. https://doi.org/10.1111/modl.12313

Sasayama, S., Malicka, A., & Norris, J. M. (2015). Primary challenges in cognitive task complexity research: Results of a comprehensive research synthesis. Paper presented in the 6th Biennial International Conference on Task-Based Language Teaching (TBLT), Leuven, Belgium.

Schoor, C., Bannert, M., & Brünken, R. (2012). Role of dual task design when measuring cognitive load during multimedia learning. *Educational Technology Research and Development*, *60*(5), 753–768. https://doi.org/10.1007/s11423-012-9251-8

Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, *58*(4), 835–874. https://doi.org/10.1111/j.1467-9922.2008.00480.x

Skehan, P. (1996a). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17*(1), 38–62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (1996b). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 17–30). Oxford, UK: Heinemann.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*(1), 1–14. https://doi.org/10.1017/S026144480200188X

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.

https://doi.org/10.1093/applin/amp047

Skehan, P. (2013). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 167–185). New York, NY: Routledge.

Skehan, P. (2014a). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam, Netherlands: Benjamins.

Skehan, P. (Ed.). (2014b). *Processing perspectives on task performance*. Amsterdam, Netherlands: Benjamins.

Skehan, P. (2015). Limited attention capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–156). Amsterdam, Netherlands: Benjamins.

Skehan, P. (2018). *Second language task-based performance: Theory, research, assessment*. New York, NY: Routledge.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*(3), 185–211. https://doi.org/10.1177/136216889700100302

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*(1), 93–120. https://doi.org/10.1111/1467-9922.00071

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). Cambridge, UK: Cambridge University Press.

Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 187–210). Amsterdam, Netherlands: Benjamins.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, *58*(2), 439–473. https://doi.org/10.1111/j.1467-9922.2008.00446.x

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Amsterdam, Netherlands: Benjamins.

Teng, L. S., & Zhang, L. J. (2016). Fostering strategic learning: The development and validation of the Writing Strategies for Motivational Regulation Questionnaire (WSMRQ). *The Asia-Pacific Education Researcher*, *25*(1), 123–134. https://doi.org/10.1007/s40299-015-0243-4

Teng, L. S., & Zhang, L. J. (2018). Effects of motivational regulation strategies on writing performance: A mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacognition and Learning*, *13*(2), 213–240. https://doi.org/10.1007/S11409-017-9171-4

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), 1–16. https://doi.org/10.1016/j.jslw.2019.100701

Tobias, S. (1986). Anxiety and cognitive processing of instruction. In R. Schwarzer (Ed.), *Self-related cognition in anxiety and motivation* (pp. 35–54). Hillsdale, NJ: Erlbaum.

Trebits, A. (2016). Sources of individual differences in L2 narrative production: The contribution of input, processing, and output anxiety. *Applied Linguistics*, *37*(2), 155–174. https://doi.org/10.1093/applin/amu006

Wang, L., & Zhang, L. J. (2019). Peter Skehan's influence in research on task difficulty: A bibliometric analysis using CiteSpace. In Z. E. Wen & M. J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 183-196). Philadelphia, PA: Benjamins. https://doi.org/10.1075/tblt.13.09wan

Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective: Influences on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 155–185). Amsterdam, Netherlands: Benjamins.

Wen, Z. E., & Ahmadian, M. J. (Eds.). (2019). *Researching L2 task performance and pedagogy: In honour of Peter Skehan*. Amsterdam, Netherlands: Benjamins.

Wickens, C. (2007). Attention to the second language. *International Review of Applied Linguistics*, *45*(3), 177–191. https://doi.org/10.1515/iral.2007.008

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing : Measures of fluency, accuracy, & complexity*. Honolulu : Second Language Teaching & Curriculum Center, University of Hawaii.

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, *39*(4), 510–522. https://doi.org/10.1016/j.system.2011.10.017

Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language Learning*, *58*(1), 151–183. https://doi.org/10.1111/j.1467-9922.2007.00437.x

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, *28*, 53–67. https://doi.org/10.1016/j.jslw.2015.02.002

Yoon, H. J. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, *66*, 130–141. https://doi.org/10.1016/j.system.2017.03.007

Zhang, L. J. (2000). Uncovering Chinese ESL students' reading anxiety in a study-abroad context. *Asia Pacific Journal of Language in Education*, *3*(2), 31–56.

Zhang, L. J. (2001). Exploring variability in language anxiety: Two groups of PRC students learning ESL in Singapore. *RELC Journal*, *32*(1), 73–91.

https://doi.org/10.1177/003368820103200105

Zhang, L. J. (2013). Second language writing as and for second language learning. *Journal of Second Language Writing, 22*(4), 446–447. https://doi.org/10.1016/j.jslw.2013.08.010

Zhang, L. J. (2016). Reflections on the pedagogical imports of western practices for professionalizing ESL/EFL writing and writing-teacher education. *Australian Review of Applied Linguistics, 39*(3), 203-232. https://doi.org/10.1075/aral.39.3.01zha

Zhang, L. J., Wang, L, & Wu, H (2017). A meta-analysis of linguistic complexity research (1990-2015) conducted within cognitive linguistics frameworks. *Fudan Forum on Foreign Languages and Literature, 10*(1), 50-62.

Zhang, L. J., & Zhang, D. (2020). Think-aloud protocols. In J. McKinley, H. Rose (Eds.) *The Routledge Handbook of Research Methods in Applied Linguistics* (pp. 301-312). London: Routledge. https://doi.org/10.4324/9780367824471-26

Appendices

Appendix A: Second Language Writing Anxiety Inventory

We would like to invite you to participate in a survey, conducted as a part of PhD project at the University of Auckland, New Zealand. This survey is designed to evaluate your second language writing anxiety when you are required to write a composition in English. This is not a test so there are no right or wrong answers. We are only interested in your personal opinion. The result of this survey will be used for research purpose so your sincere answers to these questions ensure the success of this research project. Your cooperation is highly appreciated.

Part one

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

Code: ___________ Gender: Male □ Female □ Age: ___________

Length of English learning: ________ year(s) since ______ (age)

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

Yes □ No □

If Yes, how long? _______

For what? Travel □ Study □ Other (Please specify) ________

Part two

Second Language Writing Anxiety Inventory

In this part, we would like you to help us by answering the following questions concerning your feelings of second language writing anxiety. Please give your answers sincerely, as only this will guarantee the success of the investigation. Thanks you very much for your help.

In the following section, we would like you to tell us how much you agree or disagree with the following statements by simply ticking (√) a number from 1 to 5. We are interested in your real situation and attitudes. Please do not leave out any of the items.

Totally disagree	Disagree	Partly disagree	Not sure	Partly agree	Agree	Totally agree
1	2	3	4	5	6	7

Example	I like writing in English.	1 2 3 4√ 5 6 7
1	My thoughts become jumbled when I write English compositions under time constraint.	1 2 3 4 5 6 7
2	I often feel panic when I write English compositions under time constraint.	1 2 3 4 5 6 7
3	I tremble or perspire when I write English compositions under time pressure.	1 2 3 4 5 6 7
4	I feel my heart pounding when I write English compositions under time constraint.	1 2 3 4 5 6 7
5	I usually feel my whole body rigid and tense when I write English compositions.	1 2 3 4 5 6 7
6	I freeze up when unexpectedly asked to write English compositions.	1 2 3 4 5 6 7
7	My mind often goes blank when I start to work on an English compositions.	1 2 3 4 5 6 7
8	I would do my best to excuse myself if asked to write English compositions.	1 2 3 4 5 6 7
9	Whenever possible, I would use English to write compositions.	1 2 3 4 5 6 7
10	I usually seek every possible chance to write English compositions outside of class.	1 2 3 4 5 6 7
11	I often choose to write down my thoughts in English.	1 2 3 4 5 6 7

12	I usually do my best to avoid write English compositions.	1 2 3 4 5 6 7
13	Unless I have no chance, I would not use English to write compositions.	1 2 3 4 5 6 7
14	I do my best to avoid situations in which I have to write in English.	1 2 3 4 5 6 7
15	I don't worry at all about what other people would think of my English compositions.	1 2 3 4 5 6 7
16	I'm not afraid at all that my English compositions would be rated as very poor.	1 2 3 4 5 6 7
17	I don't worry that my English compositions are a lot worse than others'.	1 2 3 4 5 6 7
18	I'm afraid that the other students would deride my English compositions if they read it.	1 2 3 4 5 6 7
19	I'm afraid of my English compositions being chosen as a sample for discussion in class.	1 2 3 4 5 6 7
20	While writing in English, I'm not nervous at all.	1 2 3 4 5 6 7
21	If my English composition is to be evaluated, I would worry about getting a very poor grade.	1 2 3 4 5 6 7
22	While writing English compositions, I feel worried and uneasy if I know they will be evaluated.	1 2 3 4 5 6 7

Great! You have completed Part Two.

Thank you for your cooperation!

Appendix B: Task Difficulty and Mental Effort Questionnaire

In the following section, we would like you to tell us how you feel about the writing task by simply ticking (√) a number from 1 to 9. We are interested in your real situation and attitudes.

1. This task required no mental effort at all.	1	2	3	4	5	6	7	8	9	This task required extreme mental effort.
2. During this task I felt extremely nervous.	1	2	3	4	5	6	7	8	9	During this task, I felt extremely relaxed.
3. This task was extremely easy.	1	2	3	4	5	6	7	8	9	This task was extremely difficult.
4. On this task I did not do well at all.	1	2	3	4	5	6	7	8	9	On this task I did extremely well.
5. This task was extremely interesting.	1	2	3	4	5	6	7	8	9	This task was not interesting at all.
6. I do not want to do more tasks like this.	1	2	3	4	5	6	7	8	9	I want to do many more tasks like this.

Great! You have completed the questionnaire.

Thank you for your cooperation!

Appendix C: Writing Tasks under the Three Planning Conditions

Task 1 (simple task[25]), pre-task planning condition

Imagine you can choose two best roommates from the four candidates when the new semester starts (two students in each dormitory). Four properties of candidates: hobbies, personality, studying style and sleeping patterns are listed in the table as follows.

Read the descriptions carefully and make a considered choice of the best roommates. Write an argumentative essay. Please provide at least 3 reasons to make your decision clear and convincing.

You have 15 minutes for planning and 25 minutes for writing. You can plan the content, organisation, and language, but your notes will be taken away before commencing of writing. Please write at least 250 words.

The best roommates you choose: _________ & __________

Name / Property	Nannan	Beibei	Niuniu	Dongdong
Hobbies	Watching movies; Listening to music; Reading books	Listening to music; Blogging; Playing the guitar	Photography; Playing football; Traveling	Listening to music; Playing football; Playing computer games
Personality	Introvert; Good at listening to others; Not good at communicating	Outgoing; Positive and enthusiastic; Happy to help others; Sometimes, ignoring others' feelings unconsciously	Like playing with others; Not good at listening to others	Outgoing; Humorous; Weak self-discipline
Sleeping pattern	Early to bed, early to rise	Late to bed, early to rise	Early to bed, late to rise	Late to bed, late to rise
Studying style	He/she likes to study in a quiet place on his/her personal.	He/she likes to read out when he/she is studying.	He/she likes to study with others.	He/she likes to study on his/her personal.

[25] The words of "simple task" or "complex task" were not shown in students' writing prompts.

Task 2 (complex task), pre-task planning condition

Imagine you can choose four best roommates from the six candidates when the new semester starts (two students in each dormitory). Six properties of candidates: hobbies, personality, studying style, sleeping patterns, favourite subjects and individual sanitary habits are listed in the table as follows.

Read the descriptions carefully and make a considered choice of the best roommates. Write an argumentative essay. Please provide at least 3 reasons to make your decision clear and convincing.

You have 15 minutes for planning and 25 minutes for writing. You can plan the content, organisation, and language, but your notes will be taken away before commencing of writing. Please write at least 250 words.

The best roommates you choose: __________ & ___________,

___________ & ___________

Name / Property	Nannan	Beibei	Niuniu	Dongdong	Kangkang	Xuxu
Hobbies	Watching movies; Listening to music; Reading books	Listening to music; Blogging; Playing the guitar	Photography; Playing football; Traveling	Listening to music; Playing football; Playing computer games	Playing the piano; Listen to music; Watching movies	Playing football; Playing computer games; Playing magic
Personality	Introvert; Good at listening to others; Not good at communicating	Outgoing; Positive and enthusiastic; Happy to help others; Sometimes, ignoring others' feelings unconsciously	Like playing with others; Not good at listening to others	Outgoing; Humorous; Weak self-discipline	Shy; Like to play with others; Weak self-discipline	Like to communicate with others and also good at listening to others; Outgoing, but sometimes kind of noisy
Sleeping pattern	Early to bed, early to rise	Late to bed, early to rise	Early to bed, late to rise	Late to bed, late to rise	Early to bed, late to rise	Late to bed, early to rise
Studying style	He/she likes to study in a quiet place on his/her personal.	He/she likes to read out when he/she is studying.	He/she likes to study with others.	He/she likes to study on his/her personal.	He/she likes to study in a quiet place on his/her personal.	He/she likes to read out in a low voice, when he/she is studying.
Favourite subjects	English, Chinese	Math, Music	Sports, Arts	Computers, Music	Music English	Math, Computer
Individual sanitary habits	Tidy, but not good at clean	Good at clean	Neutral	Good at clean	Not good at clean	Neutral

Task 1 (simple task), online planning condition

Imagine you can choose two best roommates from the four candidates when the new semester starts. Four properties of candidates: hobbies, personality, studying style and sleeping patterns are listed in the table as follows.

Read the descriptions carefully and make a considered choice of the best roommates. Write an argumentative essay. Please provide at least 3 reasons to make your decision clear and convincing.

You have no time for pre-task planning but 40 minutes for writing and online planning. You should plan while you are writing. Please write at least 250 words.

The best roommates you choose: __________ & ___________

TABLE (same as the table of PTP)

Task 1 (simple task), comparison group

Imagine you can choose two best roommates from the four candidates when the new semester starts. Four properties of candidates: hobbies, personality, studying style and sleeping patterns are listed in the table as follows.

Read the descriptions carefully and make a considered choice of the best roommates. Write an argumentative essay. Please provide at least 3 reasons to make your decision clear and convincing.

You have 40 minutes on this task. Please write at least 250 words.

The best roommates you choose: __________ & ___________

TABLE (same as the table of PTP)

Appendix D: Jacobs et al.'s (1981) Scoring Profile

ESL COMPOSITION PROFILE

STUDENT DATE TOPIC

SCORE	LEVEL	CRITERIA	COMMENTS
CONTEI	30-27	EXCELLENT TO VERY GOOD: knowledgeable • substantive • thorough development of thesis • relevant to assigned topic	
	26-22	GOOD TO AVERAGE: some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail	
	21-17	FAIR TO POOR: limited knowledge of subject • little substance • inadequate development of topic	
	16-13	VERY POOR: does not show knowledge of subject • non-substantive • not pertinent • OR not enough to evaluate	
ORGANIZATION	20-18	EXCELLENT TO VERY GOOD: fluent expression • ideas clearly stated/ supported • succinct • well-organized • logical sequencing • cohesive	
	17-14	GOOD TO AVERAGE: somewhat choppy • loosely organized but main ideas stand out • limited support • logical but incomplete sequencing	
	13-10	FAIR TO POOR: non-fluent • ideas confused or disconnected • lacks logical sequencing and development	
	9-7	VERY POOR: does not communicate • no organization • OR not enough to evaluate	
VOCABULARY	20-18	EXCELLENT TO VERY GOOD: sophisticated range • effective word/ idiom choice and usage • word form mastery • appropriate register	
	17-14	GOOD TO AVERAGE: adequate range • occasional errors of word/idiom form, choice, usage *but meaning not obscured*	
	13-10	FAIR TO POOR: limited range • frequent errors of word/idiom form, choice, usage • *meaning confused or obscured*	
	9-7	VERY POOR: essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate	
LANGUAGE USE	25-22	EXCELLENT TO VERY GOOD: effective complex constructions • few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions	
	21-18	GOOD TO AVERAGE: effective but simple constructions • minor problems in complex constructions • several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions *but meaning seldom obscured*	
	17-11	FAIR TO POOR: major problems in simple/complex constructions • frequent errors of negation, agreement, tense, number, word order/ function, articles, pronouns, prepositions and/or fragments, run-ons, deletions • *meaning confused or obscured*	
	10-5	VERY POOR: virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate	
MECHANICS	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paragraphing	
	4	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paragraphing *but meaning not obscured*	
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing • poor handwriting • *meaning confused or obscured*	
	2	VERY POOR: no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paragraphing • handwriting illegible • OR not enough to evaluate	

TOTAL SCORE READER COMMENTS

Appendix E: Questions of Semi-structured Interview

General questions:

1. Q: Do you feel anxious when you write an English composition? Why?
2. Q: Which task did you find more difficult, Task 1 or Task 2? Why?
3. Q: Which task made you more anxious, Task 1 or Task 2? Why?

Specific questions for CG:

1. What did you focus on, when you were writing the two compositions, such as content, organisation, language?
2. What were your writing habits, when you wrote the two compositions?—did you do some pre-task planning or you write as you think?

- For writing immediately (write as think): Did you often get stuck, when you were writing (you write as you think)? If yes, did you feel more anxious when you were stuck? How did anxiety influence your writing?
- For pre-task planning: What did you plan? What did you think of planning—was it beneficial or detrimental to your writing and anxious feelings?

3. Did you find whether the 40-minute time limit was enough or not?

Specific questions for OLP:

1. You were asked to write immediately when you finished reading the prompts. In that way, what was your focus/ what were you thinking, when you were writing?
2. What are your writing habits for your daily writing, pre-task planning or writing immediately?
3. What did you think of writing immediately/online planning—was beneficial or detrimental to your writing and anxious feelings?
4. Which one do you think is beneficial to your writing and anxious feelings, online planning or your normal writing habit (do some pre-task planning)?

Specific questions for PTP:

1. You were asked to do the pre-task planning for 15 minutes. What did you plan during the pre-task planning time?

2. When you wrote, did you stop writing and think about how to continue to write, such as content, organisation and language, after the 15-min pre-task planning? Or what was your focus when you were writing?
3. Did you find pre-task planning was beneficial to your writing and anxious feelings?
4. What are your writing habits for your daily writing, pre-task planning or writing immediately? Which one is better?