Copyright Statement

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

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

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

General copyright and disclaimer

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

Two studies were conducted to investigate the role of task complexity and affective factors, and task complexity and pre-task planning conditions, in L2 writing production. Two writing tasks with varying degrees of complexity in relation to the reasoning demands and the number of elements were designed and validated. Upper-intermediate English language learners were invited to participate in this study on a voluntary basis. In Study 1, one group of the learners \((n = 60)\) performed a simple and a complex version of an argumentative writing task, and their writing production was measured in terms of syntactic complexity, accuracy, lexical complexity, and fluency (CALF), as well as the organisation, content, and overall written text quality. In Study 2, one group of the learners \((n = 40)\) performed the simple and complex tasks under 10-minute pre-task planning and another group \((n = 40)\) under no-pre-task planning conditions. The learners’ writing production was compared using the same measures employed in Study 1.

In Study 1, increasing task complexity led to a significant desirable change in one dimension of both syntactic and lexical complexity, had a significant adverse effect on accuracy and fluency, and led to the enhancement of L2 content, organisation, and writing quality. As regards the mediating role of motivation and anxiety, significant moderate negative associations were found between some dimensions of maladaptive approaches to learning and some measures of L2 writing production in the complex writing task performance. Conversely, significant moderate positive correlations were found between some aspects of adaptive approaches to learning and some measures of L2 writing production in the complex L2 writing task performance.
In Study 2, a similar trend was found for the effect of increasing task complexity along the level of reasoning and the number of elements as in Study 1. The impact of 10-minute pre-task planning on L2 writing production was significantly favourable for one dimension of syntactic complexity and fluency. There were no significant effects on accuracy and lexical complexity. Nonetheless, significant positive changes for content, organisation, and writing quality were found. The findings lend partial support to the Cognition Hypothesis and the Trade-Off Hypothesis; other theoretical, methodological, and pedagogical implications are also discussed.
Dedication

To:

Shervin
Acknowledgement

I would like to express my sincere gratitude to my supervisors, Professor Lawrence Zhang and Professor Judy Parr, for their dedicated supervision, unwavering support, and insightful and intellectually engaging feedback, and to my advisor, Professor Jack C. Richards, for his wisdom, encouragement, and support of my project, and also to my family for trusting me all the way through my PhD journey.

I would also like to appreciate Professor Janet S. Gaffney for her role as my co-supervisor during the first year of my PhD research and for her feedback and encouragement, and Senior Lecturer Dr Natsuko Shintani for her feedback and evaluation of my research proposal as thorough, exceptional, and high quality. Last but not least, my special thanks go to the participants of this study. All these people have been instrumental to the successful completion of this project.
Table of Contents

Abstract ............................................................................................................................... i
Dedication .......................................................................................................................... iii
Acknowledgement .......................................................................................................... iv
Table of Contents .......................................................................................................... v
List of Tables ................................................................................................................ xxi
List of Figures ............................................................................................................... xvi
List of Abbreviations .................................................................................................. xvi
Summary of the Thesis ................................................................................................. xvii

Chapter 1. Introduction ................................................................................................. 1
  1.1. Chapter Overview ................................................................................................. 1
  1.2. Significance of Writing Skill ................................................................................ 1
  1.3. Challenges in Developing Writing Skill .............................................................. 3
  1.4. Instructional Approaches to L2 Writing ............................................................. 4
  1.5. The Role of Tasks in Writing Instructional Approaches ...................................... 6
  1.6. Proposals for Task-based Syllabus Design ......................................................... 9
  1.7. Significance of this PhD Study .......................................................................... 19

Chapter 2. Theoretical Framework and Review of Prior Research ....................... 27
  2.1. Chapter Overview ................................................................................................. 27
  2.2. Introduction ........................................................................................................... 27
  2.3. Theoretical Models of Writing ............................................................................ 29
    2.3.1. Flower and Hayes’ Writing Model (1980, 1981) ........................................... 29
    2.3.2. Hayes and Nash (1996) and Haye’s (2012) Revised Writing Model ............... 32
    2.3.3. Bereiter and Scardamalia’s (1987) Writing Models ........................................ 33
2.3.5. Galbraith’s (1999) Knowledge Constituting Model ..........37
2.3.6. Sasaki’s (2002) English as a Foreign Language Writing
   Model .....................................................................................38
2.4. Tasks in Second Language Acquisition ..................................39
2.5. Criteria for Task Difficulty and Sequencing ..............................43
2.6. The Trade-Off Hypothesis.......................................................43
2.7. The Cognition Hypothesis..........................................................51
2.8. Empirical Studies into Task Complexity and L2 Writing
   Production .....................................................................................55
2.9. Significant Prior Research into Task Complexity in L2 and EFL
   Writing........................................................................................56
2.9.1. Prior Research into Task Complexity along the Resource-
   directing Dimension..............................................................58
2.9.2. Studies into Task Complexity along the Resource-
   dispersing Dimension ............................................................64
2.9.3. Studies into Task Complexity and Second Language
   Writing Assessment ..............................................................70
2.9.4. Studies into Simultaneous Effect of Task Complexity along
   the Resource-directing and the Resource-dispersing
   Dimensions ............................................................................71
2.10. Writing Motivational Beliefs, Writing Anxiety, and Writing
   Task Complexity ...........................................................................73
2.10.1. Motivational Beliefs and Writing Task Complexity ..........74
   2.10.1.1. Writing Self-efficacy Beliefs......................................75
   2.10.1.2. Writing Task Value ....................................................76
   2.10.1.3. Writing Success Attribution .......................................77
   2.10.1.4. Writing Goal Orientation ............................................78
2.10.2. Anxiety and Writing Task Complexity ......................... 79
2.11. Summary of the Literature Review ..............................................82

Chapter 3. Research Methodology and Design .................................84
  3.1. Chapter Overview.................................................................84
  3.2. Methodology...........................................................................84
  3.3. Scope of the Research ...........................................................85
3.4. Research Objectives ................................................................. 86
3.5. Research Questions ................................................................. 87
3.6. Research Hypotheses ............................................................... 88
3.7. Pilot Study ................................................................................. 89
  3.7.1. Piloting the Questionnaires ............................................... 90
  3.7.2. Piloting the Writing Tasks ............................................... 91
  3.7.3. Piloting Data Coding and Scoring ..................................... 93
3.8. Participants .............................................................................. 93
3.9. Context of the Study ............................................................... 95
3.10. Procedure .............................................................................. 96
3.11. Materials ................................................................................ 101
  3.11.1. Tasks ............................................................................... 101
  3.11.2. Motivational Beliefs Questionnaire ................................ 104
  3.11.3. Writing Task Difficulty Questionnaire ......................... 105
  3.11.4. Writing Anxiety Questionnaire ...................................... 106
  3.11.5. Jacob et al.’s (1981) Analytical Rating Scheme ............. 107
3.12. Independent Variables .......................................................... 108
3.13. Dependent Variables ............................................................. 108
3.14. Writing Production Measures ............................................... 110
  3.14.1. Syntactic Complexity ..................................................... 110
  3.14.2. Accuracy ...................................................................... 111
  3.14.3. Lexical Diversity and Academic Vocabulary Use .......... 112
  3.14.4. Fluency ........................................................................ 112
3.15. Reliability of Data Coding and Rating .................................. 113
3.16. Statistical Analyses ............................................................... 113

Chapter 4. Results for Study 1 ...................................................... 116
  4.1. Chapter Overview ................................................................. 116
  4.2. Exploring Assumptions for Statistical Tests ....................... 116
  4.3. Task Complexity Effects on L2 Syntactic Complexity ........ 117
  4.4. Task Complexity Effects on Accuracy ............................... 118
4.5. Task Complexity Effects on Lexical Diversity and Academic Vocabulary Use ................................................................. 119
4.6. Effects of Task Complexity on Fluency ................................................... 121
4.7. Task complexity Effects on Content, Organisation, and Overall Writing Quality ............................................................................ 122
4.8. Effects of Task Complexity on Task Difficulty ........................................ 122
4.9. Modulating role of Affective Variables in L2 Writing .......... 124
  4.9.1. Modulating Role of Writing Anxiety .................................. 124
  4.9.2. Modulating Role of Writing Motivational Beliefs .............. 125
4.10. Summary of the findings for Study 1 ............................................ 127
  4.10.1. Effects on Dimensions of Syntactic Complexity .............. 127
  4.10.2. Effects on Accuracy ......................................................... 128
  4.10.3. Effects on Lexical Complexity ........................................ 128
  4.10.4. Effects on Fluency ............................................................ 129
  4.10.5. Effects on Content .......................................................... 129
  4.10.6. Effects on Organisation .................................................... 129
  4.10.7. Effects on Text Quality .................................................... 130
  4.10.8. Effects on Task Difficulty .............................................. 130

Chapter 5. Results for Study 2................................................................ 131
5.1. Chapter Overview .............................................................................. 131
5.2. Effects of Task Complexity and Planning Conditions on Subordinate Clauses per Clause ......................................................... 132
5.3. Effects of Task Complexity and Pre-task Planning Conditions on Mean Length of Clauses ......................................................... 135
5.4. Effects of Task Complexity and Pre-task Planning Conditions on Phrasal Coordination ................................................................. 139
5.5. Effects of Task Complexity and Planning Conditions on the Number of Errors per T-unit ............................................................... 142
5.6. Effects of Task Complexity and Planning Conditions on Error-free T-units per T-unit ................................................................. 146
5.7. Effects of Task Complexity and Planning Conditions on Lexical Diversity ................................................................. 150
5.8. Effects of Task Complexity and Planning Conditions on Academic Word Use .................................................. 155
5.9. Effects of Task Complexity and Planning Conditions on the Number of Words per Minute as a Measure of Fluency .......... 158
5.10. Effects of Task Complexity and Pre-task Planning Conditions on Content of L2 Writing ............................................. 162
5.11. Effects of Task Complexity and Pre-task Planning Conditions on L2 Writing Organisation ........................................... 166
5.12. Effects of Task Complexity and Pre-task Planning Conditions on L2 Writing Quality ............................................... 170
5.13. Summary of the Findings for Study 2 ........................................... 174
  5.13.1. Effects on Syntactic Complexity ........................................ 174
  5.13.2. Effects on Accuracy .......................................................... 174
  5.13.3. Effects on Lexical Diversity .............................................. 175
  5.13.4. Effects on Fluency ............................................................. 175
  5.13.5. Effects on Content ............................................................. 176
  5.13.6. Effects on Organisation ..................................................... 176
  5.13.7. Effects on Text Quality ..................................................... 176

Chapter 6. Discussion of Main Findings.............................................. 177
  6.1. Chapter Overview ..................................................................... 177
  6.2. Effects on Syntactic Complexity ............................................. 177
  6.3. Effects on Accuracy .................................................................. 183
  6.4. Effects on Lexical Complexity ................................................. 189
  6.5. Effects on Fluency ................................................................. 191
  6.6. Effects on Content ................................................................. 193
  6.7. Effects on Organisation ........................................................... 194
  6.8. Effects on Text Quality ........................................................... 195
  6.9. Effects of Task Complexity on L2 Writing Task Difficulty ....... 199
  6.10. Summary of the Discussion .................................................... 201
Chapter 7. Conclusions, Implications, Limitations, and Recommendations for Further Research

7.1. Overview

7.2. Summary of the Rationales for the Study

7.3. Summary of Main Findings

7.4. Theoretical Implications

7.5. Methodological Implications

7.6. Implications for Writing Pedagogy and Assessment

7.7. Limitations

7.8. Recommendations for Further Research

References

Appendices

Appendix A: Writing Tasks

Appendix B: Teachers’ Judgment Questionnaire

Appendix C: Pre-writing Test

Appendix D: Ethics Forms
List of Tables

Table 3.1. Descriptive Statistics for Three Groups’ Performance on Pre-writing and Placement Tests .......................................................... 97

Table 3.2. Questions on Motivational Beliefs Questionnaire .................. 105

Table 4.1. Normality Test Results .......................................................... 117

Table 4.2. Descriptive Statistics for Measures of Writing Complexity on the Simple versus Complex Tasks ........................................... 118

Table 4.3. Descriptive Statistics for Measures of Writing Accuracy on the Simple versus Complex Tasks .............................................. 119

Table 4.4. Descriptive Statistics for Measures of Writing Lexical Diversity and Academic Vocabulary Use on the Simple versus Complex Tasks ....................................................................................... 120

Table 4.5. Descriptive Statistics for Measure of Writing Fluency on the Simple versus Complex Tasks .................................................. 121

Table 4.6. Measures of Content, Organisation, and Overall Quality on the Simple versus Complex Tasks .................................................. 122

Table 4.7. Students’ Ratings of the Difficulty of the Simple versus Complex Tasks ............................................................................... 124

Table 4.8. Correlations of Dimensions of Writing Anxiety with Measures of Writing Complexity on the Simple versus Complex Tasks ....................................................................................... 125

Table 4.9. Correlations of Dimensions of Writing Anxiety with Measures of Accuracy, Lexical Diversity, Academic Vocabulary Use, and Fluency on the Simple versus Complex Tasks .............. 125

Table 4.10. Correlations of Dimensions of Writing Anxiety with Content, Organisation, and Overall quality on the Simple versus Complex Tasks ............................................................................... 125

Table 4.11. Correlations of Dimensions of Writing Motivation with Measures of Writing Complexity on the Simple versus Complex Tasks ....................................................................................... 126
Table 4.12. Correlations of Dimensions of Writing Motivation with Measures of Accuracy, Lexical Diversity, Academic Vocabulary Use, and Fluency on the Simple versus Complex Tasks ................................................................. 126

Table 4.13. Correlations of Dimensions of Writing Motivation with Content, Organisation, and Overall Quality on the Simple versus Complex Tasks ................................................................. 127

Table 5.1. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for SUB ..................................................................................................................... 133

Table 5.2. Normality Test Results for Paired Samples $t$ Tests for SUB ..................................................................................................................... 133

Table 5.3. Paired Samples $t$ Tests Results for SUB ................................................................. 134

Table 5.4. Normality Test Results for Independent Samples $t$ Tests for SUB ..................................................................................................................... 134

Table 5.5. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for MLC ........................................................................................................ 136

Table 5.6. Normality Test Results for Paired Samples $t$ Tests for MLC ................................................................. 137

Table 5.7. Paired Samples $t$ Tests Results for MLC ........................................................................................................ 137

Table 5.8. Normality Test Results for Independent Samples $t$ Tests for MLC ........................................................................................................ 138

Table 5.9. Mann-Whitney $U$ Tests Results for MLC ........................................................................................................ 138

Table 5.10. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for COR ........................................................................................................ 140

Table 5.11. Normality Test Results for Paired Samples $t$ Tests for COR ........................................................................................................ 140

Table 5.12. Wilcoxon Signed Rank Test and Mann-Whitney $U$ Tests Results for COR ........................................................................................................ 141

Table 5.13. Normality Test Results for Independent Samples $t$ Tests for COR ........................................................................................................ 141

Table 5.14. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for EPTU ........................................................................................................ 143
Table 5.15. Normality Test Results for Paired Samples $t$ Tests for EPTU ................................................................................................ 144

Table 5.16. Normality Test Results for Independent Samples $t$ Tests for ERPT ........................................................................................................... 145

Table 5.17. Paired and Independent Samples $t$ Tests Results for ERPT .. 145

Table 5.18. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for EFTPT ............................................................... 147

Table 5.19. Normality Test Results for Paired Samples $t$ Tests for EFTPT .................................................................................................................. 148

Table 5.20. Paired Samples $t$ Tests Results for EFTPT ......................... 148

Table 5.21. Normality Test Results for Independent Samples $t$ Tests for EFTPT ................................................................................................. 149

Table 5.22. Mann-Whitney $U$ Tests Results for EFTPT ......................... 149

Table 5.23. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for LD...................................................................... 151

Table 5.24. Normality Test Results for Paired Samples $t$ Tests for LD .. 152

Table 5.25. Wilcoxon Signed Rank Test Results for LD ......................... 153

Table 5.26. Normality Test Results for Independent Samples $t$ Tests for LD ................................................................................................. 153

Table 5.27. Independent Samples $t$ Tests Results for LD....................... 154

Table 5.28. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for AWU ................................................................. 155

Table 5.29. Normality Test Results for Paired Samples $t$ Tests for AWU156

Table 5.30. Wilcoxon Signed Rank Test and Mann-Whitney $U$ Tests Results for AWU ................................................................. 156

Table 5.31. Normality Test Results for Independent Samples $t$ Tests for AWU ................................................................................................. 157

Table 5.32. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for FL ...................................................................................... 159
Table 5.33. Normality Test Results for Paired Samples $t$ Tests for FL ..... 160
Table 5.34. Normality Test Results for Independent Samples $t$ Tests for FL ........................................................................................................ 161
Table 5.35. Paired and Independent Samples $t$ Tests Results for FL ...... 161
Table 5.36. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for CN ................................................................. 163
Table 5.37. Normality Test Results for Paired Samples $t$ Tests for CN ... 163
Table 5.38. Paired Samples $t$ Tests Results for CN .............................. 164
Table 5.39. Normality Test Results for Independent Samples $t$ Tests for CN .......................................................................................... 164
Table 5.40. Independent Samples $t$ Tests Results for CN ................. 165
Table 5.41. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for WORG ............................................................... 166
Table 5.42. Normality Test Results for Paired Samples $t$ Tests for WORG .......................................................................................... 168
Table 5.43. Normality Test Results for Independent Samples $t$ Tests for WORG .................................................................................... 169
Table 5.44. Paired and Independent Samples $t$ Tests Results for WORG 169
Table 5.45. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for WQ ................................................................. 171
Table 5.46. Normality Test Results for Paired Samples $t$ Tests for WQ .. 172
Table 5.47. Normality Test Results for Independent Samples $t$ Tests for WQ .................................................................................... 173
Table 5.48. Wilcoxon Signed Rank Test and Mann-Whitney $U$ Tests Results for WQ ........................................................................... 173
**List of Figures**

**Figure 3.1.** Design and Procedure for Study 1 ............................................ 99

**Figure 3.2.** Design and Procedure for Study 2 .......................................... 100

**Figure 5.1.** Bar chart with learners’ SUB means and 95% confidence intervals across four conditions .............................................. 135

**Figure 5.2.** Bar chart with learners’ MLC means and 95% confidence intervals across four conditions .............................................. 139

**Figure 5.3.** Bar chart with learners’ COR means and 95% confidence intervals across four conditions .............................................. 142

**Figure 5.4.** Bar chart with ERPT means and 95% confidence intervals across four conditions ............................................................. 146

**Figure 5.5.** Bar chart with EFTPT means and 95% confidence intervals across four conditions ............................................................. 150

**Figure 5.6.** Bar chart with learners’ LD means and 95% confidence intervals across four conditions .............................................. 154

**Figure 5.7.** Bar chart with learners’ AWU means and 95% confidence intervals across four conditions .............................................. 158

**Figure 5.8.** Bar chart with learners’ FL means and 95% confidence intervals across four conditions .............................................. 162

**Figure 5.9.** Bar chart with learners’ CN means and 95% confidence intervals across four conditions .............................................. 165

**Figure 5.10.** Bar chart with learners’ WORG means and 95% confidence intervals across four conditions .............................................. 170

**Figure 5.11.** Bar chart with learners’ WQ means and 95% confidence intervals across four conditions .............................................. 174
List of Abbreviations

CH: Cognition Hypothesis
TCF: Triadic Componential Framework
TOH: Trade-Off Hypothesis
FL: Foreign Language
L2: Second Language
CAFL: Complexity, Accuracy, Fluency, and Lexical Diversity
ILDs: Individual Learner Differences
EFL: English as a Foreign Language
TBCW: Task-based Collaborative Writing
LRE: Language Related Episodes
TBFLW: Task-based Foreign Language Writing
L1: First Language
CF: Corrective feedback
TBLT: Task-based Language Teaching
WM: Working Memory
SLA: Second Language Acquisition
SL: Second language
ESL: English as a Second Language
SLWAI: Second Language Writing Anxiety Inventory
TEFL: Teaching English as a Foreign Language
IELTS: International English Language Testing System
TOEFL: Test of English as a Foreign Language
TTR: Type-Token Ratio
Summary of the Thesis

Tasks have a central role in L2 learning and L2 writing pedagogy due to the potential transferability of task-based language learning to real-life task performance. Research has explored the impact of task type, task design and implementation features, on L2 oral language production in order to provide insights into task-based language teaching (TBLT). However, studies probing the role of task type, task design and implementation features, and the mediating role of affective factors in L2 writing are few. To fill these lacunae, this thesis extends this line of research by examining the effect of cognitive task complexity (along the degree of reasoning and number of elements) and pre-task planning conditions (presence vs. absence) on, and the mediating role of affective factors (motivational beliefs and anxiety) in, second/foreign language (L2/FL) writing production. To this end, this study draws mainly on Robinson’s Cognition Hypothesis (CH), operationalised in his Triadic Componential Framework (TCF).

Two studies investigated the isolated effect of task complexity, and the synergistic impact of task complexity and pre-task planning conditions, on L2 writing production. In Study 1, one group of learners \((n = 55)\) performed a simple and a complex version of the task, and their writing production on the tasks was measured in terms of syntactic complexity, accuracy, lexical complexity, and fluency (CALF), and organisation (cohesion and coherence), content, and overall written text quality. In Study 2, two groups of learners performed the simple and complex version of the task under 10-minute pre-task planning \((n = 40)\) versus no-pre-task planning \((n = 40)\)
conditions. The learners’ writing production across four conditions was compared using the same measures that were employed in Study 1.

The participants were recruited from among language learners taking upper-intermediate English language courses in private language schools in Esfahan, Iran. They were invited to write two essays. The two writing tasks had different degrees of cognitive complexity in relation to the reasoning demands and the number of elements, and pre-task planning conditions. The Oxford Quick Placement test version 2 and a pre-writing test were used to assess the learners’ proficiency level and writing ability, respectively. Learners whose placement scores fell into the upper-intermediate range (40-47 out of 60) and achieved good to average writing ability scores (68-85 out of 100) were recruited. Data from the pilot study, and Study 1 and 2, confirmed the validity of the researcher’s operationalisation of the simple and complex tasks employed in this study.

Study 1 addressed the isolated impact of task complexity along the resource-directing dimension (i.e., the degree of reasoning and the number of elements) on, and the mediating role of motivation and anxiety in, L2 writing production. For study 1, two types of data were collected and analysed: (a) Within group analyses by task complexity to examine the role of task complexity along the resource-directing dimension in L2 writing production, and (b) correlational analyses to examine the extent to which each dimension of motivational beliefs and anxiety relates to each measure of L2 writing production in the simple versus the complex task performance.
Study 2 inquired into the synergistic impact of task complexity along the resource-directing and resource-dispersing dimensions on L2 writing production. Within and between groups analyses were performed to examine the simultaneous effects of task complexity in terms of resource-directing (i.e., increasing the degree of reasoning and the number of elements) and resource-dispersing dimensions (i.e., +/- planning) on L2 writing production.

The results of Study 1 showed that increasing task complexity effected a significant desirable change in one dimension of syntactic and one dimension of lexical complexity, had a significant adverse effect on accuracy and fluency, and led to the enhancement of L2 content, organisation, and writing quality. As regards the mediating role of motivation and anxiety, significant moderate negative associations were found between some dimensions of maladaptive approaches to learning and some measures of L2 writing production in the complex writing task performance. Conversely, significant moderate positive correlations were found between some aspects of adaptive approaches to learning and some measures of L2 writing production in the complex L2 writing task performance.

In Study 2, a similar trend was found for the effects on CALF of increasing task complexity along the level of reasoning and the number of elements as in Study 1. With regard to the impact of 10-minute pre-task planning on L2 writing production, significant favourable results were obtained for one dimension of syntactic complexity and fluency, no effects for accuracy and lexical complexity, and significant positive changes for content, organisation, and writing quality.
The study has theoretical, methodological, and pedagogical implications. Findings of this thesis contribute to theory as they lend support to the attentional funnelling theory, with partial support to the Cognition Hypothesis (CH) and to the Trade-Off Hypothesis (TOH). The necessity of identifying the sources of syntactic complexity and accuracy or inaccuracy in L2 production was also highlighted. This identification can help in interpreting the findings as evidence for or against the TOH and has implications for task-based pedagogy and assessment. Methodologically, learners’ self-rating and learners and teachers’ judgments were confirmed as reliable methods for validating the criteria for task complexity. Moreover, the necessity of employing multidimensional measures of motivation and anxiety and developmentally sensitive, multidimensional measures of L2 writing production was established.

Regarding contributions to writing pedagogy, writing teachers may consult the TCF to adjust the complexity of the writing tasks according to the learners’ second language proficiency. Teachers can provide opportunities for using more subordination in L2 writing production and improving lexical complexity, content, organisation, and writing quality by increasing task complexity with respect to the level of reasoning and number of elements and by providing the learners with pre-task planning in performing such tasks. Teachers should promote adaptive approaches to learning via various pedagogical and assessment approaches to help learners initiate and sustain engagement with cognitively demanding tasks of L2 writing. Additionally, language test developers may well consider respective parts of the TCF in developing parallel writing assessment tasks and aligning the complexity of writing assessment tasks with the learners’ proficiency level. Test designers should develop writing tasks of
sufficient cognitive complexity and provide the learners with sufficient pre-task planning time in order to elicit their best writing performance.

More broadly, the findings support the holistic approach of TBLT in L2 writing pedagogy. Increasing writing task complexity can trigger learning processes by helping learners to notice the gaps in their interlanguage and the need for learning and development and also can afford opportunities for task-based pedagogical praxis. However, it is more likely that increasing cognitive complexity of L2 writing tasks per se leads to the increased contextualised use and potential consolidation of previously partially acquired L2 writing syntactic forms and/or more use of automatised syntactic forms that are the distinctive syntactic features of L2 learners’ proficiency level rather than stretching L2 writing syntactic complexity. The same seems to be true for the effects of providing pre-task planning and for the synergistic effects of pre-task planning and increasing cognitive complexity on L2 writing syntactic complexity.
Chapter 1.

Introduction

1.1. Chapter Overview

This chapter commences with acknowledging the significance of writing, followed by describing the emerging state of research into writing in English as a foreign/second language (L2) relative to the extensive body of research on listening, speaking, and reading skills in the field of second and foreign language acquisition research. Then, the demanding nature of writing in an L2 is underscored, and approaches to meet the challenges of L2 writing and the role of writing tasks in L2 writing instruction are discussed. Further, the importance of research into all dimensions of L2 writing, specifically the effects of writing task design and implementation features, and learner-internal factors on L2 writing production and processes is emphasised. Next, the significance of the study with regard to its role in enhancing our understanding of the impact of writing task complexity and pre-task planning conditions on L2 writing production, and the potential contributions of the study to developing research-based principles for task design, selection, and sequencing, and task-based writing assessment are discussed. Finally, the outline of the thesis is presented.

1.2. Significance of Writing Skill

Advanced writing skills can contribute to one winning access to high quality education, a high status career, and high achievements in thinking and learning.
Indeed, the significance of high achievement in writing is widely acknowledged. Graham and Perin (2007) maintain that good writing skill is typically necessary for employment in business settings and workplaces, as employees might be expected to communicate effectively and efficiently by writing e-mails, technical reports, project presentations, and other documents and, in learning, students with good writing skills can express the acquired content knowledge across the curriculum, voice their views effectively, and improve their reasoning skills relative to less skilled writers.

Likewise, developing the ability to write competently is a prerequisite for academic achievement in both native (L1) and second (L2) languages. The significance of writing in obtaining access to higher education is evidenced by writing being typically employed as one of the assessment criteria for both entry to universities and success in the university education. Writing is required for the successful completion of degree programmes in the form of writing academic reports, articles, and theses or dissertations. In addition, the importance of writing in English is further highlighted by the huge number of international students studying in English speaking countries and by the growth of English-medium universities across the globe (Tardy, 2011).

Additionally, research in L1 writing has shown how writing can enhance content learning through the role it plays in expressing, clarifying, and organising ideas (Graham & Perin, 2007; Langer & Applebee, 1985). Graham, MacArthur, and Fitzgerald (2013) have also argued that writing provides a mechanism for personal linking and influencing others, that writing is a vital tool for learning and communicating, that learners comprehend materials better if they write about them, and that teaching writing can advance learners’ reading skills.
Moreover, writing in an L2 can enhance learners’ capability to think explicitly about cognitive processes of organising and expressing their ideas (Kern, 2000). Writing can contribute to learning in planning, writing, and revising text to meet the expected writing conventions (Klein, 1999). Task-based collaborative writing (TBCW) can stretch L2 writers’ interlanguage and writing skills through incidental attention to form, termed focus-on-form episodes (Ellis, Basturkmen, & Loewen, 2001) or language related episodes (LREs) (Swain & Lapkin, 1998), and through metatalk (Ellis, 2003; Storch, 2008) about the language, writing processes, writing conventions, and expected discourse features. Task-based L2 writing can play a pivotal role in the development of L2 learners’ overall interlanguage by providing opportunities for testing hypotheses about language and writing, and for processing language through form-meaning mapping and LREs arising during L2 writing processes, and also via post-task feedback provided on the learners’ written text. Manchón (2011) has rightly highlighted the potential role of engagement in writing activities in contributing to learning to write and writing to learn. More recently, Gilabert, Manchón, and Vasylets (2016) have presented a more comprehensive picture of the potential role of writing in L2 learning and development (see also Zhang, 2013).

### 1.3. Challenges in Developing Writing Skill

Writing, whether in an L1 or an L2, is a challenging and complicated cognitive activity, posing linguistic and other cognitive demands on learners (Hayes, 2012; Kroll, 1990; Schoonen et al., 2003). Absence of immediate interlocutors (Cumming, 1990) makes writing a solitary journey (McCutchen, 1996; Kellogg, 1988; Torrance, Fidalgo, & Gracia, 2007). The need for simultaneous and continuous orchestration of
writing processes and subprocesses (Berninger, 1999) has been compared to being an operator of a complex switchboard who is attempting to balance divergent demands and constraints (Flower & Hayes, 1980). Cognitive and linguistic demands make writing an arduous task for both skilled and novice writers in any language.

In addition to the demands that writing processes impose on the cognitive resources of L1 writers, writing in an L2 can impose further linguistic demands on L2 writers’ cognition, depending on their level of L2 language proficiency and L1 background. L2 writers are “doubly challenged” (Barks & Watts, 2001, p. 250), being obliged to manage both the linguistic demands of writing in L2 and the demands of observing unfamiliar standards and conventions of writing in the target language. In fact, it can be argued that writing in an L2 can be more than a ‘triply’ challenging task, as L2 writers typically have to cope with the demands of writing in a new language, new standards and conventions of writing in the target language, and writing processes; that is, at the very least, three sources of demands compete for their cognitive resources.

1.4. Instructional Approaches to L2 Writing

There are several instructional approaches to teaching L2 writing. Ramies (1991) grouped dominant writing approaches around four foci: Form, writer, content, and reader, which I will describe briefly below. First, Form-focused pedagogy perceives writing mastery in terms of successfully acquiring the grammatical and rhetorical structures that constitute the finished text. Writing teachers adopting this approach employ various types of exercises and tasks that draw learners’ attention to surface-level textual features. Exercises and tasks in sentence combining, paragraph
formation, analysing and imitating model essays, among others, are carried out to improve stylistic proficiency. Second, a writer-focused approach, contrary to a product-focused approach, conceives of writing development as mastery of writing processes that is achieved by focusing on analysing writing task demands, planning the macrostructure, outlining, generating ideas, brainstorming, free-writing, drafting, and revising. Third approach, highly favoured in academic writing programmes, is content-focused approach. Linking writing to the specific discipline of the learners, advocates of this perspective integrate reading and writing to provide relevant knowledge and information for writing tasks. Fourth is reader-focused approach; proponents of the reader-focused approach identify genre conventions and rhetorical features of the discourse communities that learners are expected to join. Writing teachers, however, rarely base their pedagogical practices on one focus or another; they typically incorporate components from all other foci, depending on their theories of writing development and their cognition of the best practices of teaching writing (Canagarajah, 2002).

In a similar vein, Hyland (2003) identified six foci which guide writing instruction programmes: language structure, text functions, themes or topics, creative expression, composing processes, and content and genre. Hyland (2003) and Canagarajah (2002) maintain that teachers rarely base their pedagogical practice on one of the writing approaches. Instead, based on their philosophies of writing, constraints of resources in their teaching institutions, and learners’ needs, teachers adapt an eclectic multi-component approach to teaching writing. Indeed, Hyland has stressed the rarity of pedagogical practices that follow one particular focus. In a detailed discussion of the
six foci, Hyland (2003, pp. 3-22) proposed a synthesis approach that integrates process, purpose, and context perspectives into writing pedagogy:

Writing is a sociocognitive activity which involves skills in planning and drafting as well as knowledge of language, context, and audiences. An effective methodology for L2 writing teaching should therefore incorporate and extend the insights of the mains orientation in the following ways: (a) Broaden formal and functional orientations to include the social purposes behind forms; (b) Locate the process concepts of strategy, schema, and metacognition in social contexts; (c) Respect students’ needs for relevant content through stimulating readings and source materials; (d) Support genre pedagogies with strategies for planning, drafting, and revising texts; and (e) Situate writing in conception of audience and link it to broader social structures. In practice this means a synthesis to ensure that learners have an adequate understanding of the processes of text creation; the purposes of writing and how to express these in effective ways through formal and rhetorical text choices; and the contexts within which texts are composed and read which gives them the meaning (Hyland, 2003, pp. 23-24).

1.5. The Role of Tasks in Writing Instructional Approaches

Hyland argued that “tasks are the heart of a teaching unit” (Hyland, 2003, p. 112) in L2/FL writing. He underscores the significance of investigating the role of writing task types and sequencing of writing tasks in supporting and improving L2/FL learners’ writing. In his understanding, tasks are granted a central role in language teaching (see also Ellis, 2003; Nunan, 1989) and occupy a central position in language learning curriculum design. Hyland has grouped writing tasks into real-world and pedagogic tasks. Real-world tasks refer to the target tasks that learners are expected to undertake to fulfil real communicative goals such as writing an academic article for publication in a refereed journal. Pedagogic tasks are designed to improve linguistic knowledge, genre knowledge, and composing skills in order to bridge the learners’ current proficiency with the target competencies. He has argued that sequencing the writing tasks to form a coherent writing programme is imperative.
Given the significance of tasks in both task-supported and task-based language teaching (TBLT; see Ellis, 2003) and L2/FL writing instruction (Hyland, 2003), several definitions of a task (e.g., Breen, 1989; Bygate, Skehan, & Swain, 2001; Ellis, 2003; Nunan, 1989; Prabhu, 1987; Skehan, 1996) and criteria for sequencing tasks have been proposed (e.g., Breen, 2001; Long, 2015; Robinson, 2010; Skehan, 2014). For example Ellis (2003) has defined tasks in this way:

A task is a workplan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears a resemblance, direct or indirect, to the way language is used in the real world. Like other language activities, a task can engage productive or receptive, and oral or written skills, and also various cognitive processes (Ellis, 2003, p. 16).

Breen (1997, 2009) differentiated between task-as-workplan and task-as-process. He maintained that predesigned task-as-workplan becomes task-as-process due to the learners’ interpretation of the task and task demands, and the goals that the learners set to achieve. He observed that the outcome of a task is the result of the interaction between the task, learners, and task situation, and more importantly, of learners’ personal interpretation of this interaction. Ellis (2012), however, posits that Breen’s distinction of task-as-workplan and task-as-process has been overstated as, he argues, task design and implementation features can be used to influence task performance, although Ellis also calls for inquiries into how learner factors influence task performance.

In writing, Manchón (2014) has proposed that conceptualisation of tasks in SLA be expanded to be applicable to writing due to writing specific learner-internal factors
(e.g., task representation and goal setting). She foregrounds learner-internal factors (e.g., learners’ interpretation of writing tasks) which is more aligned with Breen’s (1997, 2009) definition of task-as-process and sociocultural perspectives of TBLT. Despite evident differences among these perspectives, the common theme is the appreciation of the role of tasks, task conditions, and learner factors in task performance, and ultimately, learning.

Obviously, conceptualising and proposing notional definitions of tasks in L2 writing and subsequently operationalising these definitions and subjecting them to empirical tests to examine the effects of task inherent, task condition, and learner-internal, factors on task performance is important. However, Long (2015, 2016) posits that the real and unresolved issues, which have implications for both task-based language teaching and assessment, comprise identifying transferable and generalisable features of the task types and developing task difficulty and sequencing criteria. By identifying transferable features of the task types, a reasonable number of tasks can be designed that can be taught over a reasonable timescale in a language teaching curriculum. And by developing task sequencing criteria, these tasks can be sequenced and implemented in a comprehensive L2 learning programme. Additionally, identifying generalisable features of the task types can enhance the feasibility and validity of task-based assessment.

Studies into the issue of generalisable features of tasks and transfer of task-based language learning are scant (e.g., Benson, 2015), and Long (2015, 2016) has called for further inquiries into this highly significant area. However, with regard to task sequencing, several proposals exist in the literature, which aim at facilitating the
planning side of TBLT and designing a coherent task-based syllabus. Research is needed to examine the predictive power of these proposals to provide insights into developing principles for task-based syllabus design.

1.6. Proposals for Task-based Syllabus Design

In the following section, the proposals for task-based syllabus design are presented briefly. The comprehensive discussion of the proposals that this study draws on is reserved for the Literature Review chapter.

Candlin (1984) proposed that a syllabus be negotiated between the teacher and learners and be selected “from among a bank of items and procedures” (p. 36), and Breen (2001) suggested that tasks are about meaning rather than forms and recommended grading tasks according to the logic of sequencing tasks to solve a series of connected problems. Others have suggested sequencing tasks from simple to complex based on their proposed criteria or framework of task difficulty or complexity. In 1987, Prabhu proposed sequencing tasks based on “reasonable challenge” (p. 55) as defined by Vygotsky’s (1978, cited in Prabhu, 1987, p. 86) “zone of proximal development” (ZDP). ZDP refers to the distance between an individual’s independent problem solving capability and his or her problem solving ability under assistance from a more capable individual.

Long and Crookes (1992) and recently Long (2015, 2016) have proposed using a pre-determined syllabus with tasks as units for organising and sequencing the syllabus. They highlighted the importance of conducting a need analysis to identify target tasks that students will need to do in real-life situations, developing task types from target
tasks, and finally designing and sequencing pedagogic tasks based on the complexity of the tasks to form a coherent task-based syllabus. Long and Crookes (1992) did not recommend an explicit taxonomy for sequencing tasks; however, their conceptualisation of task complexity and factors leading to the complexity of tasks can be gleaned from their suggestions:

The number of steps involved, the number of solutions to a problem, the number of parties involved and the saliency of their distinguishing features, the location (or not) of a task in displaced time and space, the amount and kind of language required, the number of sources competing for attention, and other aspects of the intellectual challenge a pedagogic task poses… (p. 44).

Long and Crookes also recommended taking into account the role of pedagogic options that are used to implement tasks in tasks sequencing (e.g., open versus closed, planned versus unplanned, whole-class versus small-group interactions) to modulate task complexity.

Skehan (1996, 1998, 2002, 2003, 2009, 2014) and Skehan and Foster (2001) have argued for a balanced interlanguage growth and posited that in TBLT putting too much emphasis on meaning at the expense of form can lead to developing communicative strategies and employing lexicalised language in successful meaning making without paying due attention to form that can have negative consequences for interlanguage development. Skehan (1996) has proposed a framework for accomplishing a balanced growth of accuracy, fluency, and complexity. He has highlighted the significance of both task sequencing (syllabus) and task implementation (methodology) in achieving this goal.

Skehan (1996) proposed criteria for establishing task difficulty, including task complexity (syntactic and lexical difficulty), cognitive complexity (content required
for successful task performance), cognitive processing (amount of online mental processing needed during the task performance), cognitive familiarity (does the task refer to ready-made solutions?), time pressure (how quickly should the task be done?), modality (speaking vs. writing), scale (number of participants), stakes (how important is it to do the task?), and control (can participants exert an influence on how a task is done? and can participants negotiate the task goals?). Skehan, following Willis (1996), also stressed the importance of three phases of methodology (pre-task, during task, and post-task activities) in successful implementation of TBLT.

Nunan (1989, 2004) has proposed that tasks can be sequenced according to the cognitive and performance demands they impose on learners, suggesting sequencing tasks from comprehension-based activities through controlled production to real-world tasks that engage learners in communication.

Ellis (1987, 2003) has recommended that two parallel strands of tasks (unfocused and focused) be employed in task-based and task-supported language teaching. In an unfocused task, meaning is primary and learners can employ their own linguistic resources to complete the task. Focused tasks can be used to extend the learners’ linguistic resources; still meaning is the focus of the task, but form (e.g., grammar, vocabulary, and/or pragmatics) is discussed by the participants in task performance. In task-supported language teaching the curriculum is sequenced, based on linguistic items and tasks are used to provide situational practice of the linguistic/notional/functional items that are presented and practiced prior to their production in the final stage of instruction. In contrast, in task-based language teaching, tasks are the units of the syllabus and language learning arises incidentally
from tasks performance and from pre-task, during task, and post-task exploitations (Skehan, 2014, 2016). Contrary to Ellis (2003), who recommends that both task-supported and task-based language teaching be used in language teaching, Long (2015, 2016), Robinson (2011), and Skehan (2014) reject task-supported language teaching. It is in task-based language teaching that developing criteria for task difficulty and sequencing is an unresolved issue (Long, 2015, 2016).

Ellis (2003) has proposed four criteria (i.e., Input, Conditions, Processes, and Outcomes) for task difficulty. Input is the information that is needed for the task completion, conditions refer to the conditions under which a task is performed (e.g., monologic vs. dialogic), processes refer to cognitive processes involved (e.g., inferencing), and outcome refers to whether the task output is simple or complex. These four criteria can be manipulated by task designers to form tasks of varying degrees of difficulty. Like Long (2015, 2016), Long and Crookes (1992), Robinson (2011) and Skehan (1998, 2014, 2016), Ellis (2003, 2013) has also proposed that methodological options (e.g., pre-task planning time), which are used to implement task-based instruction, can be employed to further modulate the complexity/difficulty of tasks to adjust them with the learners’ proficiency level. Despite some differences in their proposals, these scholars have recommend that tasks be designed according to task difficulty/complexity criteria and sequenced from easy to difficult to form a task-based syllabus, and that methodological options be used to further modulate the task complexity/difficulty and implement the task-based syllabus.

Framework (TCF) for task design and classification and a Simple/Stabilising interlanguage. Automatising access to interlanguage, Restructuring and Complexifying of interlanguage (SSARC) model for task sequencing (Robinson, 2010). Robinson differentiates task complexity from task difficulty defining the former as:

…the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner. These differences in information processing demands, resulting from design characteristics, are relatively fixed and invariant (Robinson, 2001a, p. 29).

The TCF defines task complexity along two main dimensions: the resource-directing and resource-dispersing dimensions. The resource-directing refers to the conceptual demands that tasks impose on the learners, including the tasks’ reference to past or present events, here and now, few or many elements, and more or fewer reasoning demands. The resource-dispersing dimension, including availability or unavailability of pre-task planning time and prior knowledge of the task, and the number of tasks to be completed imposes procedural demands on the learners’ cognitive resources.

Robinson (2010, 2011a, 2011b) posits that tasks should be sequenced only based on the cognitive demands of the tasks, and task complexity in task-based syllabus should be increased along the resource-dispersing dimensions first, and then along the resource-directing dimensions, and finally along both dimensions to reach maximum complexity, so that tasks approximate real-life tasks in order to promote language development through his proposed SSARC model for task sequencing. He has proposed that, in the first instance, tasks be simple in both dimensions (e.g., + planning time and – reasoning) to draw upon learners’ stable interlanguage. In the second step, complexity should be increased by removing planning time to elicit
speedier production to improve automatisation. Finally, complexity should be increased along the *resource-directing* dimension to approximate target tasks’ complexity level, by adding reasoning for instance, to destabilise and restructure learners’ current interlanguage system.

As follows from the preceding review, proposals for sequencing tasks in TBLT have evolved, with understandable overlaps among the proposals. Each framework is an expansion of the previous one in terms of both conceptualising task complexity or difficulty and detailed suggestions or frameworks for task sequencing. Researchers have carried out studies to investigate the premise of these criteria or frameworks for task sequencing. These criteria and corresponding research have concentrated mostly on oral communication (e.g., Ellis, 1987, 2003, 2009a; Nunan, 1989, 2004; Long & Crookes, 1992; Skehan, 1996, 1998, 2002, 2003; Robinson, 2001a, 2001b, 2005, 2007a, 2010).

However, conceptualisation of, and inquiries into, task design and sequencing in L2 writing are scarce, although tasks play a central role in L2 writing classes and the significance of developing pedagogical tasks to form comprehensive writing programmes is underscored by (Hyland, 2003). As such, conceptualisation of, and research into, L2 writing tasks complexity or difficulty, sequencing, and implementation are necessary for L2 writing pedagogy. One such conceptualisation of using tasks in writing instruction is through task-based collaborative writing instruction (TBCWI).
In line with Long’s (2015) model of TBLT, comprehensive TBCWI programmes can be formed by using writing tasks as units of syllabus. Writing task types can be identified by a needs analysis and pedagogical writing tasks can be designed and sequenced based on validated criteria of L2 writing task complexity. Writing task performance can afford opportunities for learners to consolidate and/or expand their interlanguage system and writing skills. These opportunities can arise from writing task performance and/or language and writing learning episodes (LWLEs) prior to, during, and after collaborative writing task performance. LWLEs can arise from writing task performance in and of itself and/or be due to problems arising from breakdowns in meaning making and/or extending meaning. LWLEs refer to an incidental focus on form and writing skills in meaning-focused writing tasks, the purpose of which is negotiating meaning and achieving the objectives of the writing tasks. In TBCWI, in addition to writing to learn the language (Manchón, 2011; e.g., vocabulary, grammar, and pragmatics) through writing task performance and/or LWLEs, learning different dimensions of writing processes (i.e., goal setting, planning, idea generating, organising, composing, and revising) can also be achieved.

TBCWI might encompass teacher-whole class collaboration and student-student collaboration that focuses on discourse features and writing processes involved in successful performance of writing tasks. TBCWI may also involve concentrating on the language via discussing what language to employ to achieve the goal of the task and on analysing model tasks to identify sentence and discourse level features. Also, in TBCWI learners’ attention can be directed to the content of the task (e.g., identifying the arguments and counterarguments to be included in the writing task). In TBCWI learners can then focus on collaborative composing of the task, followed
by revisions that can result in consolidating and/or stretching language and writing skills.

The final episode can be the individual task performance in the same task type, followed by teacher and peer feedback. In other words, writing tasks can be sequenced based on the validated taxonomy of task complexity and can be implemented by using pre-task, task, and post-task methodological options to modulate task complexity to create task-based opportunities (e.g., contextualised pedagogical instruction) for interlanguage and writing skills development. The stepping stone toward such a writing pedagogy is developing criteria for designing and sequencing such tasks. Further, developing developmentally sensitive measures of writing and interlanguage development in writing (e.g., L2 writing syntactic complexity) are desirable in order to be able to assess the extent to which writing tasks performance and/or learning opportunities arising from task performance lead to affordances for consolidating and/or stretching learners’ interlanguage and writing skills.

Therefore, two of the most compelling endeavours for writing research are identifying developmental levels and task design and implementation features that can yield opportunities for consolidating and/or stretching L2 learners’ interlanguage system and writing skills. Additionally, to design pedagogical writing tasks validated criteria of task design and sequencing is required. Fortunately, studies have reached the point that some developmentally sensitive measures of L2 writing syntactic complexity (Norris & Ortega, 2009) have been designed and validated, and the Triadic Componential Framework (TCF; Robinson, 2007a, 2011a) is proposed as
criteria for pedagogical task design and sequencing. The next step is to design tasks based on the proposed criteria, validate the operationalisation of task complexity in these tasks, and then measure their impacts on L2 writing production. These studies should confirm or disconfirm the validity of the dimensions of such criteria and reveal the effect of performing tasks of varying degrees of complexity on L2 writing production and, ultimately, on language and writing learning and development. The results of such studies can also provide insights into whether L2 learners’ L2 processing capacity is limited, or they have access to multiple attentional resource pools. These findings can contribute to theory, research, and pedagogy. In short, the results can contribute to resolving one of the remaining issues of TBLT, that is, validating criteria for task design, sequencing, and selection in task-based language teaching and assessment.

In an attempt to contribute to resolving one of the issues of TBLT, this study endeavours to investigate the extent to which manipulating writing task complexity affects L2 writing production and affords opportunities for consolidating and/or stretching learners’ interlanguage system in terms of L2 writing syntactic complexity. This study examines the role of task design and implementation features, also referred to as task requirements (Macaro, 2014), namely, increasing task complexity in terms of the degree of reasoning and the number of elements and pre-task planning conditions, and the moderating role of affective factors, in writing production to expand the emerging body of research from a psycholinguistic perspective. Notwithstanding harsh critiques that have been levelled at this strand of inquiry (e.g., Manchón, 2014), the role of variety in tasks has been acknowledged in sociocultural theory as the variety in artefacts that are deployed to improve learning (Ellis, 2009b).
Psycholinguistic research can elucidate the role of task design and implementation features (learner-external) and affective (learner-internal) factors, among others, in writing task production and the opportunities that writing task performance afford for writing to learn (WTL) in terms of consolidating and/or expanding learners’ interlanguage system and/or expanding writing quality and skills development. Research from this perspective can also provide insight into the processing capacity of L2 learners.

Inquiries into task design and implementation features and the interaction of these features with learner-internal factors can provide potential insight into designing pedagogical tasks and sequencing them. Pedagogical tasks that provide opportunities for language practice and/or development and/or writing skills practice and development or, at the very least, prompt learners to produce L2 writing productions and create opportunities for contextualised pedagogical intervention such as written corrective feedback, conferencing, and reformulation.

One such approach to designing and sequencing pedagogical tasks involves exploring task design and implementation features and the mediating role of affective factors to pinpoint the extent to which these features impose cognitive demands on the learners and affect their writing production and, ultimately, provide opportunities for language learning and writing development. Promoting learners’ language and writing development through manipulating tasks performance might be preferable to explicit teaching of linguistic forms and writing skills, as contextualised task performance is less likely to disturb communicative purpose of tasks (Adams, Alwi, & Newton, 2015), and learning occurring in task performance and/or opportunities
arising from task performance is more likely to be transferred to real-life task performance (Long, 2015).

Although there is a scarcity of hard evidence for pedagogical interventions conducive to transfer of learning (Long, 2015), “transfer-appropriate processing” (Morris, Bransford, & Franks, 1997; cited in Long, 2015) lends support to task-based writing pedagogy; transfer of learning is more likely to occur if learning takes place in tasks involving similar processes to the tasks for which transfer is sought. In response to calls for a principled basis to task-based syllabus design and implementation (Skehan, 1996), two psycholinguistic models have hypothesised the impact of task design and implementation features and also learner factors on L2 production. In this study, the predictions of these competing hypotheses are examined to provide potential insights into developing principles for task-based language and writing pedagogy and assessment.

1.7. Significance of This PhD Study

Hyland (2003) has highlighted the significance of tasks in writing and posited that designing pedagogical writing tasks and sequencing them is imperative to improve learners’ linguistic knowledge, genre knowledge, and composing skills. Likewise, Manchón (2011) has rightly maintained that engagement in writing activities can contribute to learning to write (LW) and to other learning, including content and language knowledge (see also, Zhang, 2013). Nevertheless, inquiries into writing task design and implementation features and learner factors to provide potential insights into pedagogical task design and sequencing are few. The results of such studies can contribute to validating task complexity or difficulty criteria for task design and
sequencing and provide insights into praxis by revealing the extent to which writing
tasks performance in and of itself promotes language and writing skills consolidation
and/or development and/or, at the very least, affords opportunities for contextualised
pedagogical interventions. The results can also provide insights into whether learners
have limited or multiple attentional resources for L2 processing, which is very
important for understanding L2 performance and learning.

To fill these lacunae, an emerging body of research (Ellis & Yuan 2004; Frear &
2013, 2014; Ong & Zhang, 2010, 2013, Ruiz-Funes, 2015) has been carried out into
the impact of task design and implementation features, and also into learner-internal
factors on L2 writing tasks performance. These researchers have examined the
predictions of two frameworks of task complexity/difficulty (those of Robinson,
2001a, 2011b, and Skehan, 1998a, 2014). However, the results are far from
conclusive mainly due to variety in the design of studies and the multiplicity of
factors involved.

Although writing task types, time-on-task, and measures used in previous studies
vary considerably, which prohibits comparison of the findings; authors defend the
findings of their own studies in the face of contradictory findings of others. Drawing
conclusions and generalisations requires an extensive body of research from multiple
perspectives. For instance, the impact of planning conditions on writing production
might differ depending on the differences between the complexity threshold of the
writing tasks, operationalised as simple versus complex, task types, and learners’
motivational beliefs and anxiety, among other factors. Further, the effects of task
complexity on higher-order skills such as cohesion and coherence and task achievement (strength of arguments and rebuttals) should also be addressed (Kuiken & Vedder, 2008). Moreover, recruiting participants with similar proficiency levels and educational backgrounds, providing them with sufficient orientation about task requirements, and ensuring that participants have the requisite skills to undertake the writing tasks can enhance the design of the study.

In view of the preceding discussion, task design and implementation features, and individual learner differences should be inquired into from multiple perspectives (e.g., psycholinguistic and sociocultural) in order to elucidate the effects of these variables on learners’ interpretation of tasks, writing processes, production, and learning and development opportunities. To extend this line of research inquiry, this study draws on mainly Robinson’s Cognition Hypothesis (Robinson, 2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b), Kellogg’s (1996) working memory (WM) model in L1 writing, and Skehan and Foster’s Trade-Off Hypothesis (Skehan, 1998a, 2001, 2003, 2009, 2014; Skehan & Foster, 1999, 2001) to examine their predictive power in L2 writing task performance in order to provide potential insights into L2 and L2 writing pedagogy and assessment. Specifically, research into the effects of task complexity, task conditions, and task difficulty on complexity, accuracy, lexis, and fluency (CALF) in second/foreign language (FL/L2) writing production is inconclusive. Hence, this study is designed to investigate the role of task complexity, task planning conditions, writing motivation, and anxiety in FL/L2 writing production.
To address these research gaps, this thesis contributes to the emerging body of research into the effect of cognitive task complexity (i.e., along the degree of reasoning and the number of elements) and pre-task planning conditions (presence vs. absence) on, and the mediating role of affective factors (motivational beliefs and anxiety) in, L2/FL writing production. The study also extends prior research by employing multidimensional measures of motivation, anxiety, and developmentally sensitive, multidimensional measures of syntactic and lexical complexity. Using developmentally sensitive, multidimensional measures of L2 writing production is particularly important, as the findings can differentiate whether manipulating task design and implementation features results in the use of more forms that are the distinctive features of the learners’ current interlanguage system (e.g., more subordinate clauses for upper-intermediate proficiency level), or whether it also expands their interlanguage system by eliciting forms that are the prominent features of the advanced level of proficiency (e.g., more grammatical metaphors). To elaborate, if increasing task complexity for upper-intermediate proficiency level learners elicits more forms that are the distinctive features of the advanced level, it can be argued that increasing task complexity stretches learners’ interlanguage system. Robinson (2011b) predicts that increasing cognitive complexity of tasks along the resource-directing dimension will increase the use of developmentally more advanced forms of language.

To this end, I designed and conducted two studies to investigate the isolated effect of task complexity and the synergistic impact of task complexity and pre-task planning conditions on L2 writing production. In Study 1, one group of learners ($n = 55$) performed a simple and a complex version of the task, and their writing production
on the tasks was measured in terms of syntactic complexity, accuracy, lexical complexity, fluency (Calf), and organisation (cohesion and coherence), content, and overall written text quality. In Study 2, two groups of learners performed the simple and complex versions of the task under 10-minute pre-task planning ($n = 40$) versus no-pre-task planning ($n = 40$) conditions. The learners’ writing production across four conditions was compared using the same measures that were employed in Study 1.

The participants were recruited from among language learners taking upper-intermediate English language courses in private language schools in Esfahan. They were invited to write two essays. The two writing tasks had different degrees of cognitive complexity in relation to the reasoning demands and the number of elements, and pre-task planning conditions. The Oxford Quick Placement test version 2 and a pre-writing test were used to assess the learners’ proficiency level and writing ability respectively; learners whose placement scores fell into the upper-intermediate range (40-47 out of 60) and achieved good to average writing ability scores (68-85 out of 100) were recruited.

This study is grounded in the psycholinguistic perspective of TBLT, which mainly adopts a positivist approach to second language acquisition (SLA) research. The advocates of this perspective argue that task design and implementation features influence how learners’ attend to the intended dimensions of their production. In line with prior research along this line of inquiry, this study intends to contribute to and extend our understanding of the role of tasks in L2 writing production. Specifically, this study endeavours to seek answers to the following research questions:
1. What is the effect of task complexity (less vs. more) on writing production (i.e., CALF, content, organisation, and text quality)?

2. What is the effect of task complexity (less vs. more) on perceived task difficulty?

3. What is the relationship between learners’ writing motivational beliefs’ and each measure of writing production (i.e., CALF, content, organisation, and text quality)?

4. What is the relationship between learners’ anxiety and each measure of writing production (i.e., CALF, content, organisation, and text quality)?

5. What is the simultaneous effect of task complexity (less vs. more) and planning (present vs. absent) conditions on writing production (i.e., CALF, content, organisation, and text quality)?

The independent variables of this study, namely, task complexity along the resource-directing dimension (i.e., the degree of reasoning and the number of elements) and the resource-dispersing dimension (absence vs. presence of pre-task planning time), and the mediating variables of motivation and anxiety are defined, and the prior research pertaining to each dimension is discussed in the Literature Review chapter. Additionally, the research methodology, scope, objectives, questions, and hypotheses are presented in the Methodology chapter, as are the comprehensive definitions of the dependent variables.

1.8. Outline of the Thesis

This thesis comprises seven chapters. Chapter 1, the Introduction, sets the scene for the whole thesis by mapping the significance of L2 writing, challenges that L2
writers face, and the instructional approaches employed in L2 writing pedagogy to date. In this chapter, the role of tasks in L2 writing pedagogy, and the unresolved issue of developing validated criteria for task complexity in order to design and sequence pedagogical tasks are discussed. Subsequently, two main frameworks for task complexity are presented. Finally, an emerging body of research agenda that contributes to task-based L2 writing pedagogy are identified, and how this PhD study extends this strand of research is described.

Chapter 2 initially recaps the significance of research into the role of task design and implementation features and affective factors within the task-based research framework. Then, writing models are reviewed briefly with regard to their predictions regarding the effect of task design and implementation features on the dimensions of L2 writing production. Subsequently, Kellogg’s (1996) theoretical model of L1 writing working memory (WM), Skehan and Foster’s Trade-Off Hypothesis (Skehan, 1998a, 2001, 2003, 2009, 2014; Skehan & Foster, 1999, 2001), and Robinson’s Cognition Hypothesis (Robinson, 2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b), theories that this study draws on are presented. Finally, research literature pertaining to the variables of this study is reviewed.

Chapter 3 illustrates the methodological paradigm that this thesis is grounded in and the methods for data collection and analyses. This is followed by presenting the research objectives, questions, hypotheses, and the research design used to address the research questions. Subsequently, the pilot study, the description of the participants, context of the study, research design, data collection procedures, and materials used to collect the data are described. Finally, the independent and
dependent variables of the study, measures used for data coding and rating, and the reliability coefficients of data coding and rating are elaborated on.

The results of Study 1 and 2 are presented in chapters 4 and 5, respectively. The effects of cognitive task complexity on, and the modulating role of motivation and anxiety in, the dimensions of syntactic and lexical complexity, accuracy, fluency, content, organisation, and overall written text quality are presented in chapter 4. The chapter concludes with a summary of findings for Study 1. Chapter 5 describes the results of Study 2, the simultaneous effects of task complexity and pre-task planning on the dimensions of syntactic and lexical complexity, accuracy, fluency, content, organisation, and overall written text quality. This chapter closes with a summary of the findings for Study 2.

Chapters 6 and 7 present the Discussions and the Conclusions of the study. In Chapter 6, the results of Study 1 and 2 are interpreted in relation to research questions, the predictions of the theoretical frameworks that guided this study, and the pertinent prior research findings. In Chapter 7, the implications of the study to theory, research, and pedagogy are discussed, as are the limitations of the study and future directions for research.
Chapter 2.

Theoretical Framework and Review of Prior Research

2.1. Chapter Overview

In this chapter, first the rationales of the study are presented. This is followed by the critical review of the theoretical models of writing and their predictions about the effects of task design, implementation, and learner, factors on dimensions of writing production. Then, two competing hypotheses, the Cognition Hypothesis and the Trade-Off Hypothesis/Limited Attention Model are described. Next, the previous literature on the effect of task complexity on L2 writing production is critically reviewed. Finally, prior research into the moderating role of writing motivational beliefs and writing anxiety in L2 writing production is discussed.

2.2. Introduction

Prior research has demonstrated similarities in L1 and L2 writing cognitive processes with regard to planning (Jones & Tetroe, 1987), revising (Cumming, Rebuffot, & Ledwell, 1989), and editing (Hull, 1987), despite the obvious differences between L1 and L2 writing. For instance, studies have revealed that skilled and less skilled L2 writers are relatively similar to L1 writers in planning and revising. L1 and L2 skilled writers plan and revise more at discourse levels, whereas less skilled L1 and L2 writers tend to plan more at word and phrase levels (Ramies, 1987; De Larios, Murphy, & Manchón, 1999). Owing to these commonalities between L1 and L2 writing, L1 models lay the foundation for L2 writing models (Krapels, 1990).
However, writing models make no predictions in relation to the effects of task design, implementation, and individual learner differences, factors on the dimensions of learners’ L2 writing production (Kormos, 2011; Ong & Zhang, 2013; Ruiz-Funes, 2015; also see the critical review of L1 and L2 models in the following section). As such, a few extant prior studies (e.g., Frear & Bitchener, 2015; Ong & Zhang, 2013) have explored the impact of these factors on L2 writing production via linking L2 writing research with task-based research. In line with this strand of research, this study, placed within the theoretical framework of task-based research, endeavours to extend the emerging body of research into the role of cognitive task complexity in L2 writing production in line with the predictions of the Cognition hypothesis (Robinson, 2011a, 2011b) in order to provide insights into the role of task-based L2 writing in writing to learn (Manchón, 2014) and learning to write and potentially into task-based writing assessment.

Specifically, this thesis contributes to research into the effect of cognitive task complexity (i.e., in terms of the degree of reasoning and number of elements) and pre-task planning conditions (presence vs. absence) on, and the mediating role of affective factors (motivational beliefs and anxiety) in, second/foreign language (L2/FL) writing production. The study also extends prior research by employing multidimensional measures of motivation and anxiety, and developmentally sensitive, multidimensional measures of syntactic and lexical complexity. Using developmentally sensitive, multidimensional measures of L2 production (e.g., L2 writing syntactic complexity) is particularly important, as the findings can differentiate whether manipulating task design and implementation features affects
using forms that are the distinctive features of the learners’ current interlanguage system, or whether it also extends their interlanguage system.

However, to justify the necessity of conducting this study within task-based research, a brief critical review of writing models is presented in the following section. More specifically, L1 and L2 writing models are reviewed to elucidate the extent to which the impact of task design and implementation features, and affective factors, and the interaction effect of these factors on the dimensions of L2 learners’ writing production and text quality are realised and predicted in these writing models.

2.3. Theoretical Models of Writing

In this section, the theoretical models of L1 and L2 writing are discussed to identify the extent to which these models theorise the effects of task design and implementation features, and individual learner differences, on the syntactic complexity, accuracy, lexical complexity, and fluency of writing production.

2.3.1. Flower and Hayes’ Writing Model (1980, 1981)

Critiquing stage models of writing (Britton et al., 1975; Rohman, 1965, cited in Flower & Hayes, 1981), which depict writing in linear stages of prewriting, writing, and re-writing, as inaccurate and inadequate, Flower and Hayes (1981) argued these models do not pay attention to the cognitive processes of writing and focus on the writing product. Hence, they proposed a contrasting cognitive writing process model. The model is a classical writing model of cognitive writing processes that was developed based on data from verbal protocol analysis and text analysis of skilled and less skilled writers.
Three core components, namely, the writing task environment, the writer’s long term memory, and the underlying writing processes comprise the model. Each component consists of subcomponents. The rhetorical problem and text written so far constitute the external task environment component. Knowledge of topic, writing plans, and audience make up the writer’s long term memory component. Planning, translating, and reviewing, which function under the control of monitoring, compose the underlying writing processes component. The distinctive feature of this model is the depiction of writing as a recursive process, that is, planning, translating, and reviewing are processes that occur at all stages of the writing process, recursively rather than linearly.

Due to the hierarchical structure of this model, the components consist of subcomponents. Planning encompasses goal setting, generating ideas, and organising ideas. Translating involves transferring generated ideas into language that requires lexical and syntactic knowledge and motor skills of spelling. Lack of automaticity in translating can impose cognitive demands on writers and can interfere with other cognitive processes (Flower & Hayes, 1981). Reviewing refers to the cognitive processes of reading the produced text and evaluating it against the set-goals and revising it to achieve the required goals if necessary.

The model depicts underlying cognitive processes of skilful writers, which result in desired writing outcomes. However, no hypothesis is proposed about the roles of writing task characteristics and implementation features in the dimensions of writing performance (e.g., syntactic complexity, accuracy, & fluency) that can provide
potential insight into L2 task design and implementation in a comprehensive L2 writing programme. Understanding how automaticity in employing these cognitive processes and using advanced forms of linguistic encoding develops is highly significant for L2 pedagogy. A worthwhile endeavour would be to understand how task design and implementation features affect using advanced linguistic forms accurately, fluently, and appropriately. Flower & Hayes’ (1981) model provides no explanation of how linguistic encoding and cognitive writing skills become automatic or how teaching can optimise skill development. The model does not address questions such as: Should a teacher teach the cognitive processes in separate stages and/or should free writing be encouraged from the early stages of the writing-to-learn journey to enhance linguistic encoding and writing? Which approach will effectively and efficiently lead to automaticity in desired linguistic encoding and writing skills? What other factors such as individual difference (ID) factors and writing task design and implementation features might affect linguistic encoding, cognitive writing processes, and writing production?

These questions are not addressed by these and similar models, a lack which justifies linking L2 writing research with task-based research to address them and further enhance our understanding of the development of writing processes and linguistic encoding in L2 writing production. Specifically, no predictions are made in these and similar writing models about writing task design and implementation features and writing conditions and ID factors that might influence different aspects of L2 writing production, including syntactic complexity, accuracy, lexical complexity, and fluency (CALF). Task-complexity theory (Robinson, 2001a, 2015) from a task-based research perspective has proposed such hypotheses and provides guidance on how to
sequence tasks to promote the automaticity and development of desired linguistic encoding in L2 learning.

2.3.2. Hayes and Nash (1996) and Hayes’ (2012) Revised Writing Model

Building upon Flower and Hayes’ (1981) model, Hayes and Nash’s (1996) Revised Writing Model adds a new social environment component to the model, which consists of audience, collaborators, and individual differences. Four components, namely, cognitive processes, long-term memory, working memory, motivation, and affect compose the revised model. The cognitive processes in the revised model are organised and labelled differently from that of the Flower and Hayes’ (1981) model. Problem-solving, decision making, and inferencing comprise reflection. Planning and organising are subsumed under problem-solving and decision making, respectively. Inferencing reflects writers’ assumption about their audience’s interest, and decision-making refers to the choices the writers make with regard to their point of emphasis in organising their writing and selecting competing opinions.

In the revised model, long-term memory contains task schemas, linguistic, and genre knowledge. Task schemas help writers identify writing goals, procedures for writing, and evaluation processes. In the revised model, translation (i.e., putting thoughts into ideas) and revision are renamed as text production and text interpretation, respectively. In the revised model, rereading of the text for comprehension and text evaluation is depicted as the vital component of the text interpretation which, if inadequate, results in poor text quality (Hayes & Nash, 1996). Motivation and working memory are given a pivotal role in the revised model; it is proposed that non-automated processes impose cognitive demands on working memory, and that
the goals, predisposition, beliefs, and attitudes of writers as the components of motivation are important in writing processes.

Recently, Hayes (2012) has proposed a detailed adult writing model as work in progress which comprises three levels: control level, process level, and resource level. Motivation, goal setting, current plan, and writing schema constitute the control level, which forms and drives writing activity. Writing processes and task environment make up the process level, and the resource level comprises attention, long-term memory, working memory, and reading. Despite the advancements in the newly proposed writing model, similar to previous models no predictions are proposed regarding the role of task design and implementation features that can provide insights into task-based L2 writing pedagogy.

**2.3.3. Bereiter and Scardamalia’s (1987) Writing Models**

Bereiter and Scardamalia’s (1987) have proposed two writing models: The knowledge telling and knowledge transforming models. The former reflects writing processes of less skilful writers and the latter skilful writers’ writing processes. From a developmental perspective, writing is proposed to progress from knowledge telling for children and less skilled writers to knowledge transformation for adults and skilled writers (Bereiter, Burtis, & Scardamalia, 1988).

The Knowledge Telling model consists of sequential processes. These processes are said to be activated by mental representations of the writing task, which lead to finding topic and genre. Identification of topics and genres results in memory search which retrieves related content and discourse knowledge from long term memory.
After that, the retrieved content and discourse are evaluated; if the retrieved data are found inappropriate, the writer will repeat the cycle until appropriate data are retrieved; the appropriate data are transcribed and, finally, the writer updates the written text to reflect current mental representation. Bereiter and Scardamalia (1987) maintain that this strategy is well suited for writing, which requires sequential detailed descriptions of events, genres such as recounts and narratives.

The Knowledge Transforming model extends the Knowledge Telling model by adding problem analysis, goal setting, content problem space, and rhetorical space. By problem analysis and goal setting, the writer identifies the content, audience, and genre. This model proposes that writing is a problem-solving process involving solving two kinds of problems, namely, content space and rhetorical space problems. The former refers to content knowledge and the latter encompasses discourse knowledge. The model proposes that knowledge transforming occurs by interaction between content and rhetorical spaces. That is, problem transfers between two spaces result in changes in content and organisation of the writers’ knowledge.

With regard to revision, Scardamalia and Bereiter (1985) propose three metal processes, namely, Compare, Diagnose, and Operate (C.D.O) operations. To revise, by comparing the written text with the intended text, the writer identifies any discrepancies, which need to be attended. This is followed by diagnosing the nature of problem, which ends with correcting or generating new content strategy.

This model and previous writing and revision models, as general cognitive models do not explore elements of the environment, of which the task is one. And, as writing is
a taught skill, they give little direction as to instruction. Therefore, further research building upon previous emerging research from a task-based research framework (e.g., Ellis & Yuan 2004, Frear & Bitchener, 2015; Kormos, 2011; Kuiken & Vedder, 2007, 2008, 2011, 2012; Ong & Zhang, 2010, 2013; Ong, 2013, 2014) is warranted to extend the knowledge base on the role of task design and implementation features and ID factors in L2 writing production.


Kellogg (1996) has highlighted the significance of working-memory capacity in writing and proposed that the processes of formulation, execution, and monitoring in writing production place cognitive demands on the working memory components: The visuo-spatial sketchpad, the central executive, and the phonological loop (Baddeley, 1986). In Kellogg’s three level writing production model, formulation comprises planning and translating conceptual ideas and rhetorical goals into text; execution refers to creating the text by hand or word processing it, and monitoring to evaluating and revising the produced text by reading and editing it. For formulation, writers utilise their working memory functions to retrieve ideas and knowledge to construct a new composition (planning) and linguistic elements, namely, lexical, syntactic, and rhetorical items to express the ideas and knowledge (translating). These processes function concurrently and, depending on the writers’ skill, influence the working memory capacity. Skilled writers would need to allocate attentional resources to activate the linguistic information essential for writing performance (Kellogg, 2001), while less skilful writers would require making more demanding cognitive effort to write (Kellogg, et al., 2013; MacCutchen, 1996).
Kellogg et al.’s (2013) evaluation of Kellogg’s (1996) model based on current research has confirmed the principal assumption of the model that planning, translating, and reviewing place cognitive demands on the central executive. Also, they confirmed that linguistic encoding (translation of ideas into a sentence) draws on verbal working memory and the central executive. It was also confirmed that planning sentences that involve concrete words, but not abstract words, engage the sketchpad. It also appears that planning does not engage the spatial subcomponent of the sketchpad. Furthermore, the assumption that editing places no cognitive demands on the phonological loop was seriously challenged.

One of the main predictions of Kellogg’s (1996) model is the limited capacity of central execution. Therefore, depending on the task type, task design and implementation features, and task performance conditions, writers might prioritise and focus on one writing process over the other and one dimension of linguistic encoding over the other (syntactic complexity vs. accuracy). To elaborate, one factor that might affect writing production is the availability of planning time that might lead to attentional funnelling. For instance, the availability of pre-task planning might funnel the attention to idea generation during pre-task planning and free attention to linguistic encoding during translating (Kellogg, et al., 2013). In contrast, when a learner writes under time constraints, the learner may have to share limited working memory capacity between the planning (generating ideas) and translating (linguistic encoding), as Ellis (2005, 2009b) has proposed. This implies that in less cognitively complex tasks, writers’ familiarity with task types, and the availability of planning time may afford funnelling attention to formulation, execution, and reviewing by
imposing less cognitive demands on writers’ limited working memory capacity, which, in turn, can lead to improved linguistic production. Similar to other models of writing production, this model does not provide specific hypothesis on the role of task design and implementation factors on the specific dimensions of the L2 writing production (i.e., CALF). However, task-based cognitive models of L2 performance draw on WM models to provide specific predictions about the role of task design and implementation factors on the specific dimensions of the L2 writing production. Therefore, Kellogg’s model is relevant to this thesis, particularly Kellogg, et al.’s (2013) attentional funnelling theory and their proposal that different dimensions of writing production might draw on different components of WM.

2.3.5. Galbraith’s (1999) Knowledge Constituting Model

Galbraith (1999) posits that “knowledge encoded in sentences is represented, implicitly, within a distributed network of conceptual relationships, and that ideas are synthesised by constraint satisfaction within this network, rather than being directly retrieved” (p. 141). Galbraith argues that content generation in writing is influenced by the range of units activated within the network, the complexity of links among the activated units, writers’ linguistic knowledge and goals, the form of output, and types of planning. The model predicts that the dispositional dialectic process, which is very similar to the free-writing strategy (Elbow, 1973) and associate writing (Bereiter, 1980), leads to generating new ideas in the translation process.

Some other models consider planning responsible for idea generation and translating for linguistic encoding. Galbraith’s (1999) prediction is contrary to the widely accepted role of planning in writing. The role of pre-task planning versus no-pre-
task-planning conditions in writing production, including content, organisation, and overall writing quality are investigated in this PhD study.

2.3.6. Sasaki’s (2002) English as a Foreign Language Writing Model

Sasaki (2002) studied the cognitive writing processes of three groups of writers: Expert writers, novice writers before instruction, and novice writers after instruction. Her findings revealed that the differences between expert and novice writers lie in the planning and revising strategies they use. Expert writers used global planning and rhetorical revising. In contrast, novice writers employed thematic (rough global) planning and did not engage in rhetorical refining. Concentrating on the differences between expert and novice writers, Sasaki’s (2002) writing model distinguishes the differences in planning and revising processes of these groups and does not make any prediction with regard to the effect of task complexity and ID factors on task performance.

In view of the preceding discussions, although task environment and ID factors are embodied in the recent models of writing (e.g., Hayes, 2012), these models do not make any specific prediction regarding the role of task complexity factors in writing production and written text quality, which is of great significance for writing task selection, sequencing, and assessment. As such, in line with some prior studies in L2 writing production, which have been conducted within a task-based research framework (Ellis & Yuan 2004; Frear & Bitchener, 2015; Kormos, 2011; Kuiken & Vedder, 2007, 2008, 2011, 2012; Ong & Zhang, 2010, 2013; Ong, 2013, 2014), I examined the effect of task-complexity along the resource-directing and the resource-dispersing dimensions on L2 writing production and written text quality. A
detailed account of the theoretical basis of task-based research, sociocultural and psycholinguistics perspectives of tasks, and the role of tasks in TBLT is given in the following sections.

2.4. Tasks in Second Language Acquisition (SLA)

The central role of tasks in SLA is underlined in both sociocultural and psycholinguistic perspectives of SLA (Benson, 2015; Crookes & Gass, 1993; Ellis, 2003, 2009a; Ellis & Shintani, 2013; Long, 2015, 2016; Nunan, 1989; among others). Specifically, Hyland (2003) has maintained that writing tasks are integral components of write-to-learn and characterised tasks as the essential facet of teachers’ planning and delivery of a writing instruction. Tasks play a fundamental role in writing classes, as they influence learners’ learning experiences by triggering cognitive mental processes that the learners coordinate over all the recursive processes of writing, including planning, activation of schematic topical and linguistic knowledge, formulation of ideas, execution, and monitoring (Flower & Hayes, 1981; Flower et al., 1990; Kellogg, 1996; Kellog et al., 2013).

Tasks are ‘workplans’ that are meaning-based (language or content as the core of meaning making activity), set goals for learners to achieve, are evaluated based on task outcome, and are of real-world significance (Ellis, 2000, 2003, 2009a; Skehan, 1998a). In short, tasks refer to meaning making activities which are accomplished by using the language as a vehicle (Hyland, 2003; Ellis, 2003). Tasks can be both input-providing and output-prompting (Ellis, 2009b). Input-providing tasks can be in the form of listening and reading tasks, although speaking and writing tasks also provide input. The production of one interlocutor can take the form of input for the other.
Output-prompting tasks are presented in speaking and writing production tasks. Writing production task, for instance, performing an argumentative writing task entails meaning making (e.g., arguing for or against a cause), setting goals (e.g., supporting an argument), an outcome that can be evaluated (e.g., the produced written text), and awareness of the intended audience. These components of tasks show the value of tasks in the real-world and the necessity of attending to both meaning and language for the successful completion of tasks (Horowitz, 1986a, 1986b; Long & Crookes, 1992; Swales, 1986).

There are two main perspectives of TBLT: Sociocultural and psycholinguistic. From a psycholinguistic perspective, task performance can lead to language use and acquisition by prompting mental processes (Ellis, 2009b) such as noticing the gap and hypothesis testing in “pushed-output” (Swain, 2005). Principally, psycholinguistic models attempt to provide insights into understanding language learning and development by distinguishing the role of task types, and task design and implementation features, in language use and opportunities for the development of the different dimensions of L2 production, namely, syntactic and lexical complexity, accuracy, and fluency (CALF). The models intend to provide insights into task difficulty, selection, and sequencing, as they consider tasks units of the syllabus in TBLT.

From a sociocultural perspective, learners do not passively implement tasks as imposed by the design features of the tasks. Instead, participating learners actively contribute to and define the tasks and set goals. Learners actively co-construct the task performance environment, intended audience, objectives, the required processes,
and linguistic resources to accomplish the goals of the assigned tasks. As such, task execution is the result of the interaction between participants and task features (Ellis 2009b). That is, in addition to task design features, which potentially draw learners’ attention to different aspects of task performance, specific participants’ co-construction of a task might also trigger mental processes that could differ if the same task were to be performed by different participants.

However, as Ellis (2009b) clarifies, the role of task design and implementation features is acknowledged in sociocultural theory as the variety in the features of the artefacts employed in learning, and empirical studies employing this theory have stressed the role of task types, and task design and implementation features, in L2 production (see Skehan, 2009; Robinson, 2011b). Hence, the two perspectives are complementary and research within both the sociocultural and psycholinguistic perspectives of TBLT can surely contribute to our understanding of task-based learning.

Despite the acknowledged central role of tasks in both sociocultural and psycholinguistic perspectives of TBLT, mainly due to the transferability of task-based learning to real-life task performance (Ellis, 2013; Lightbown, 2008; Long, 2015), there are remaining issues that future research is warranted to contribute to resolving, if TBLT is to be grounded in sound research evidence rather than intuition and armchair speculations. Ellis observes that the principal problem is grading tasks based on task difficulty criteria that take into account both linguistic and cognitive difficulty and expose learners to a reasonable challenge aligned with their developmental level. Implied in Ellis’s argument are, indeed, two challenges: First,
identifying the developmental level of the learners’ language system, and, second, designing and grading tasks that are aligned with learners’ developmental level to provide affordances for consolidating and/or stretching their interlanguage system. However, as Ellis (2013) suggests, TBLT was proposed in response to the difficulty of matching instructional syllabuses to learners’ developmental readiness, what Corder (1967) termed “built-in syllabus”. TBLT affords opportunities for learners to attend to and learn the forms that are aligned with their developmental level, without imposing forms to be learned in a linear fashion.

Three main issues in TBLT in need of research evidence are: (a) Developing criteria for task difficulty and sequencing, (b) identifying underlying generalisable features of real-life task types, and (c) establishing transfer of task-based learning to real-life task performance (Long, 2015, 2016). Validated criteria for task difficulty and sequencing would make designing pedagogical tasks types and sequencing them possible. Identifying underlying generalisable features of real-life task types to develop pedagogical tasks would facilitate designing comprehensive second language learning programmes, as there might be numerous tasks under each task type category, and time constraints might not allow including all those tasks in TBLT programmes. Identifying such underlying generalisable features can contribute to the feasibility and validity of task-based assessment too. Practically, only a manageable number of assessment tasks can be included in a task-based assessment package. Establishing the transfer of task-based learning to at least similar real-life task types is highly desirable. After all, the main principle of TBLT and the philosophy in which it is grounded is its claimed conduciveness to the transfer of learning from classroom task performance to real-life task performance.
Admittedly, an extensive body of research is needed to provide the much needed research evidence for each unresolved category. Even under each category, there are large numbers of issues that each requires a body of research. As I discussed in the Introduction chapter, in response to resolving the issue of criteria for task difficulty for developing and sequencing pedagogical tasks, criteria have been proposed (e.g., Ellis, 2003; Skehan, 1996, 2009; Robinson, 2001a, 2007a, 2011b). Robinson has proposed the most comprehensive framework for such a purpose, operationalised in the TCF.

2.5. Criteria for Task Difficulty and Sequencing

Two main hypotheses have been proposed to address task complexity and sequencing in TBLT. To elaborate, Skehan (1998, 2001, 2003, 2009, 2014) and Robinson (2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b) have proposed two psycholinguistic models that address how task design features influence learning by imposing different levels of cognitive demands on learners’ cognitive resources and drawing learners’ attention to specific dimensions of language use, production, and learning. These models have been used both in studying the impact of task design features on task performance and manipulating features of instructional tasks (Robinson, 2011b). These two models are discussed in more details in the following sections, as they form the theoretical foundation of this study.

2.6. The Trade-Off Hypothesis

The Trade-Off Hypothesis (TOH)/Limited Capacity Model (LCM) claims that tasks that are more demanding use more attentional resources, so that less attention is available for the learners to focus on form. Accordingly, this model proposes sequencing tasks from least to most complex tasks to create opportunities for the learners to attend to both meaning and form. The model suggests that manipulating task features (e.g., structured vs. unstructured tasks) and task implementation (e.g., providing pre-task and/or post-task transcription) can afford opportunities for the learners to attend to different dimensions of their production, leading to the balanced development of their interlanguage system in terms of complexity, accuracy, and fluency (CAF).

The model classifies the task variables under two general groups: Task features, which refer to how information is structured and what elements are manipulated, and task implementation, which refers to pre-task activities (e.g., pre-task planning time and task repetition), task performance under varying conditions, and post-task activities. The model claims that by manipulating these task features, learners’ limited attentional resources can be drawn to a certain aspect of language production, which can promote the development of CAF under certain task design and performance conditions. Put simply, Skehan (1998a, 2001, 2003, 2009) posits that as learners have a limited attentional resource, and thus limited processing capacity, they can focus only on one aspect of language production (syntactic complexity, accuracy, or fluency) at a given time while performing a given task. In other words, Skehan argues that the manipulation of task and implementation variables affect the fluency, accuracy, and complexity of production, and as task difficulty increases, learners will attend to only one aspect of production at the expense of the others.
Specifically, the TOH/LCM proposes a framework for task difficulty and sequencing tasks from simple to complex ones. Skehan (1996) has proposed criteria for establishing task difficulty, including code complexity (syntactic and lexical difficulty and range), cognitive complexity (the difficulty of the content required for successful task performance), cognitive processing (amount of online mental processing and calculation needed during the task performance), cognitive familiarity (the degree of the participants familiarity with the task topic and procedures of the task completion; reference to ready-made schematic solutions vs. creation of the solutions), time pressure (the amount of time allocated to complete the task), modality (speaking vs. writing; listening vs. reading), scale (the number of factors related to task performance, e.g., number of participants), stakes (the importance of successful completion of the task), and control (the level of control that participants can exert on the goal, procedure, etc. in task performance). To elaborate, higher code complexity, cognitive complexity, cognitive processing, communicative stress, scale, and stakes, and lower cognitive familiarity and control make tasks more difficult. Additionally, Skehan argues that speaking and listening tasks lead to more pressure and are more difficult than writing and reading tasks. Skehan proposes that tasks be sequenced from simple to more difficult based on his task difficulty criteria to form a task-based syllabus.

Additionally, Skehan (1996) has stressed the importance of three phases of methodology (pre-task, during task, and post-task activities) in implementing a task-based syllabus. He has argued that teachers can further adjust the difficulty of tasks and draw leaners’ attention to different dimensions of L2 production (i.e., CAF) by
making implementation decisions in pre-task, during task, and post-task stages. To elaborate, in the pre-task planning phase, teachers can draw learners’ attention to complexity by giving them sufficient pre-task planning. Actions can be taken to adjust the code and cognitive complexity and cognitive processing involved in task production. During task performance, the communication stress (time, modality, scale, stakes, and control) can be adjusted to draw learners’ attention to the intended dimensions of task production. For the post-task phase, by using post-task activities such as public performance, task repetition, and similar task performance and reminding the learners that a certain activity will follow the main task performance, learners’ attention can be drawn to a certain aspect of L2 production (i.e., CALF), which can lead to restructuring, consolidating, and fluency, Skehan posits.

In short, Skehan (1996) proposes a framework for task-based syllabus design and implementation to sequence and implement the tasks in a principled way. By making task design and implementation decisions, Skehan argues, intended changes in the learners’ interlanguage system can be brought about. The underlying principle is that the learners’ attention is limited (Skehan, 1996, 2009); they can only attend to one dimension of their production (CAF) at a time. Hence, learners should be assigned tasks manipulated in terms of task difficulty and implementation in order to enhance the pre-targeted dimension of their L2 production.

Skehan (2009) has argued that his review of research into the effects of task types and task characteristics has revealed the scarcity of simultaneous improvements in accuracy and complexity. Hence, he posits, that the findings support the Trade-Off Hypothesis. He also argues that concurrent enhancements in few documented cases
are the result of easing attentional limitation. He explains these results in L2 performance by drawing on the stages of speech production in Levelt’s (1989, 1999) model of speaking. Skehan’s summary of the impact of the task type studies show that narrative task performance results in greater complexity, but lower accuracy and fluency, and personal information exchange task performance raises accuracy and fluency, but not complexity. For the effect of task characteristics on L2 performance, tasks with concrete or familiar information and tasks containing clear structure advantage accuracy and fluency; interactive tasks confer advantages to accuracy and complexity; tasks requiring information manipulation result in greater complexity; post-task conditions raise accuracy, and pre-task planning results in higher complexity and fluency, but not accuracy.

Skehan (2009) explains the simultaneous improvements in complexity and accuracy by recourse to Levelt’s (1989, 1999) model of L1 speaking and argues that the coinciding advantages for complexity and accuracy are not the result of cognitive complexity and access to multiple attention resource pools, as Robinson (2011a, 2011b) argues. Rather, Skehan posits that it is the result of easing and freeing limited attentional resources, due to particular planning opportunities, and structure and information integration that provide supportive conditions for learners to overcome normal limited attentional constraints. In interpreting Foster and Skehan’s (1999) result, Skehan argues that teacher-led planning resulted in combining planning-as-rehearsal and planning-as-complexification. The former led to increases in accuracy and the latter to greater complexity. Regarding Tavakoli and Skehan (2005) and Tavakoli and Foster’s (2008) findings, Skehan argues that support provided by the structured nature of the task resulted in accuracy and information integration in
complexity. Planning-as-rehearsal and planning-as-complexification affect different stages of L2 production based on Levelt’s model. Planning-as-rehearsal supports the Formulator and planning-as-complexification the Conceptualiser, and ease of attention resulting from these two sources of support results in joint improvements. Skehan posits that these results are due to support conditions not increasing task complexity.

To explain Foster and Skehan’s (2013) result of the decision-making task with post-task condition, in which learners produced significantly greater accuracy and complexity, Skehan argues some learners prioritised accuracy and some others complexity. As the correlation between accuracy and complexity was small and not significant, he concluded that at the individual level the trade-off effect is the norm. In fact, he posits that to establish simultaneous increases in two dimensions of production, both significant positive experimental effects and a significant positive correlation for the dimensions that increase concurrently need to be demonstrated. However, the purpose of experimental designs is to account for individual differences (Dörnyei, 2007), and the lack of a significant positive correlation between variables with significant positive experimental effects (e.g., complexity and accuracy) does not necessary confirm the trade-off effect. One way to address Skehan’s concern is to identify the number of students that experience simultaneous increases and the trade-off effect at the individual level through frequency analysis.

Additionally, Skehan (2009) argues that depending on the complexifying/pressuring and/or easing and focusing nature of tasks, tasks influence the availability of limited attention to the Conceptualiser and Formulator, which, in turn, affects CALF. To
elaborate, tasks requiring information integration call for extensive attention by the Conceptualiser to create a complex message which in turn results in complex language use. Tasks demanding less frequent lexical retrieval result in the Formulator facing difficulty with lemma retrieval that consumes attention and leads to more sophisticated lexis and lower accuracy. Tasks with macrostructure consume less attention by the Conceptualiser and release attention for the Formulator resulting in accuracy and fluency. These results are better accounted for by the Trade-Off Hypothesis, drawing on Levelt’s model, Skehan posits.

Recently, Skehan (2014) reported on the results of several recently completed research projects, which he argues, all support the trade-off interpretation. Based on prior research findings and results reported in his 2014 book, Skehan has proposed micro and macro level principles for task-based pedagogy. He has proposed micro within-lesson level principles based on task design and implementation features and macro level implementation by using project works to link a series of tasks to form task-based lessons for long timescales. Skehan has highlighted the role of the task performance in noticing and hypothesising that triggers learning new forms and that of post-task exploitation in complexifying/extending (linking the new form with the interlanguage system), in restructuring/integrating (reorganising the interlanguage system) and repertoire creation, and that of task implementation features in achieving accuracy, automaticity, and lexicalisation.

Moreover, Skehan (2014) argues that noticing is necessary but not sufficient to result in the acquisition of new forms and disagrees with the proposition that interaction contains all needed to acquire new forms. He argues that noticing the deficiencies
should come from the learners and be nurtured by the teacher. For hypothesising, the need to do the task makes language salient and prompt learners to hypothesise how language works. Teachers can follow-up to reinforce the hypothesis, correct a mistaken hypothesis, link it with other parts of language and extend it and help learners relate it to the other parts of the interlanguage and complexify their interlanguage. He also argues that teachers can help learners restructure/reorganise their interlanguage through post-task exploitation. For instance, in the case that regular and irregular past tenses have to be integrated into the interlanguage system, teachers can help learners differentiate the rule-based system (regular past tense) from the exemplar-based system (irregular past tense).

For Skehan (2009), control (accuracy and fluency) can be brought about by using task types, task characteristics, and task implementation features to draw learners’ attention to accuracy and fluency. He has proposed teachers can use traditional practices to develop control in leaners’ performance. Automatisation can arise from task performance under supportive conditions, but automatisation also involves generalisation challenge, so learners need to perform tasks in pressurised conditions to gain confidence in performing the task in a wide range of contexts, Skehan posits. The final stage of learning is lexicalisation, which, Skehan believes, can be achieved via drawing learners’ attention to the Formulator. However, Skehan argues that achieving lexicalised language needs a longer of period of exposure to language. There are several differences between forms that are acquired through task performance in task-based learning and traditional approaches. Some of these are: Forms become salient and noticed as a result of task performance, and learners more likely notice the forms that they are developmentally ready to acquire (Skehan,
2014); these acquired forms are more likely to be transferred to real-life situations (Long, 2015); forms are not prescribed by syllabus designers, and tasks are not used for the situational practice of the prescribed forms (Ellis, 2013). In view of the above, Skehan proposes criteria for task difficulty and sequencing, and principles for a task-based approach and also calls for further research into task design and implementation features in order to provide insight into task-based L2 pedagogy. Indeed, a body of research is required to examine Skehan’s criteria for task difficulty, his rationale for task sequencing, and his principles for TBLT.

2.7. The Cognition Hypothesis

Robinson (2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b) has also proposed task complexity and sequencing models for task-based L2 pedagogy. However, contrary to LCM, the Cognition Hypothesis, operationalised in the Triadic Componential Framework (Robinson, 2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b), proposes a multiple resources perspective on language processing that predicts that conceptual demands of the tasks prompt learners to draw on multiple resource pools and simultaneously process different dimensions of language. Accordingly, Robinson argues that the conceptual demands of more cognitively complex tasks draw learners’ multiple attentional resources to complexity and accuracy and facilitate language learning and development. He argues that increasing cognitive task complexity will increase the conceptual demands of the task which, in turn, will result in interlanguage development, the use of developmentally more advanced forms of language, and will also improve the accuracy of L2 production. In L2 writing, sub-clausal complexity is shown to be the distinctive feature of the advanced proficiency level and subordination the upper-intermediate proficiency level in terms of syntactic
complexity (Norris & Ortega, 2009). Hence, in L2 writing, increases in sub-clausal complexity and more accurate L2 writing production for upper-intermediate level students as a function of increasing task complexity will confirm one of Robinson’s claims.

The Triadic Componential Framework (2011b) comprises three dimensions: Task complexity, task condition, and task difficulty. Resource-directing and resource-dispersing compose task complexity. The resource-directing dimension imposes varying degrees of cognitive demands on the attentional resources of the learners. The degree depends on the tasks’ conceptual demands such as their reference to past or present events, here and now, few or many elements, and more or fewer reasoning demands. The resource-dispersing dimension imposes procedural demands on the learners’ cognitive resources and includes planning time, prior knowledge of the task, and the number of tasks to be completed, among others.

Robinson (2001a, 2011b) argues that increasing task complexity along the resource-directing dimension leads to the simultaneous enhancements of accuracy and complexity of language production, as learners have to utilise their multiple attentional resources to meet the multiple conceptual demands of the task. However, fluency of the learners’ language production will deteriorate as students have to process language. Conversely, increasing task complexity along the resource-dispersing dimension will result in decreased fluency, accuracy, and complexity of learners’ language production because it will impose procedural demands on learners’ working memory (see Robinson 2011a for the complete framework).
Robinson (2010) has proposed that tasks of varying degrees of complexity be designed and sequenced from simple to complex ones based on task complexity dimension and SSARC, respectively. He has proposed that first tasks be simple along both the resource-dispersing and the resource-directing dimensions. Then, the complexity of the tasks should be increased along the resource-dispersing dimension (e.g., no-pre-task planning) in order to promote real-time access to the existing linguistic resources. The complexity should be increased along the resource-directing dimension next to stretch, restructure and complexify the interlanguage system. Finally, the complexity should be increased in both dimensions, so that pedagogical tasks approximate real-life tasks.

Task condition consists of two sub-categories: Participation (e.g., dialogic vs. monologic) and participant (e.g., same vs. different gender). Robinson argues that task condition should be kept constant across the versions of the same task with varying degrees of complexity to ensure the transfer of the schemata developed in the simpler version to the complex version of the task. Task difficulty includes individual differences (ID) in learners’ ability (e.g., working memory capacity) and affective factors (e.g., anxiety) that result in between-learner differences. Robinson argues that ID factors cannot be employed as the principled basis for syllabus design, but the ID should be accounted for through methodological decisions such as task-ability matching and providing support based on the learners’ needs.

Like Skehan (2014), Robinson (2011b) draws on Levelt’s (1989, 1999) model of L1 speaking to provide the theoretical rationale for his prediction of task demands on L2 production. He argues that increasing cognitive task complexity leads to more effort
at the preverbal conceptualisation stage due to conceptual demands imposed on the Conceptualiser. This, in turn, results in encoding the conceptual information via linguistically appropriate forms at the lexicogrammatical stage, leading to joint enhancements in accuracy and complexity in L2 production. However, as discussed previously, Skehan argues joint improvement in accuracy and complexity is rare, as L2 learners’ processing capacity is limited. Therefore, although both Robinson and Skehan draw on the same theory to provide rationales for their claims, their interpretation and claims are contrasting. Skehan’s predictions are discussed above, and Robinson’s predictions, which my study examines, are presented below.

One main and five ancillary claims are proposed by the CH (Robinson, 2011a). The main pedagogical claim of the CH is pedagogical tasks should be sequenced from simple to complex based on the cognitive complexity criteria, operationalised in the TCF. However, research evidence is needed to establish the validity of the TCF. Révész (2014) proposes approaches to studying the validity of the framework (see also Révész, & Gurzynski-Weiss, 2016). Robinson has also identified five ancillary claims of the CH: (a) Increasing task complexity along the resource-directing dimension will advantage complexity and accuracy in monologic task performance but will only advantage accuracy in dialogic task performance and will have an adverse effect on fluency in both task types; (b) cognitively complex task performance will lead to more interaction and negotiation of meaning, and noticing and uptake of the forms that are made salient in the input; (c) complex task performance will lead to higher depth of processing and long-term retention of the input provided (e.g., conversational feedback); (d) simple to complex sequencing of tasks should result in greater automaticity relative to sequencing of the tasks in other
sequences such as complex, simple, and complex, and (e) ability and affective factors will differentiate learners’ performance more clearly in complex than simple task performance, and affective factors will be more influential in interactional rather than monologic task performance. Additionally, Robinson has proposed that increasing task complexity along the resource-directing and reducing task complexity along the resource-dispersing dimensions will have positive synergistic effects on L2 production.

In summary, Robinson (2011a, 2011ab) and Skehan (2009, 2014) have proposed criteria for task design and implementation and sequencing and both have called for further research in order to validate their criteria for task design and implementation to provide insights into developing research-based principles for task-based L2 language pedagogy. Robinson calls for research first to establish the isolated effect of task design features (simple vs. complex) on task performance and then to examine whether such an effect is multiplied by the sequence in which tasks are performed. As pointed out in previous section, a body of research is warranted to validate the proposed criteria for task complexity/difficulty, examine the rationales for sequencing tasks from simple to complex/difficult, and explore the predictive power of these proposals in L2 writing production. This study seeks to examine the predictions of the CH; as such, relevant previous studies are reviewed in the following sections.

2.8. Empirical Studies into Task Complexity and L2 Writing Production

To date, there are contradictory findings with regard to the claims of the CH (Robinson, 2011a), and the results (e.g., those of Kuiken & Vedder, 2007; Robinson,
2007) lend partial support to the predictions of the hypothesis. In line with Kormos and Trebits (2011), I would suggest that these inconsistent findings can be accounted for by the multiplicity of factors involved in language task performance. In other words, in some tasks the two dimensions of the resource-directing and the resource-dispersing may be simultaneously present, which can confound the findings. Also, as there is clear variability in the research designs of the previous studies due to the multiplicity of factors involved, a significant body of research into each factor from multiple perspectives is required to reach a valid conclusion. Further research is warranted to isolate and investigate different dimensions of task complexity and their potential impacts on task processing, production, and interlanguage development in order to provide research-based insights into building task-based language teaching and assessment principles.

2.9. Significant Prior Research into Task Complexity in L2 and EFL Writing

As research into cognitive task complexity in L2/FL writing and writing task performance is in its embryonic stage (Ong & Zhang, 2010, 2013, Ruiz-Funes, 2015), research findings on the effects of task complexity and L2 writing production are far from conclusive; this makes generalisations of findings and the forming of research-based principles for task-based syllabus design, pedagogy, and assessment within the domain of L2 and FL writing still problematic (Frear & Bitchener, 2015; Ruiz-Funes, 2015). That is, in order to provide research-based evidence, further research is required to isolate and investigate the multiple factors affecting writing production. This study is an endeavour to contribute to this emerging area of research. To date, research findings have indicated varying and contradictory impact of the cognitive task complexity on syntactic complexity, accuracy, fluency, and lexical variation
(Frear & Bitchener, 2015; Ruiz-Funes, 2015; also evidenced in the review of empirical studies in the following section).

However, as I will discuss in the concluding section of this chapter, the controversy in the findings, to date, cannot be attributed to the inadequacy of task complexity theory. This state of affairs is largely due to the paucity of research into the impact of cognitive task complexity on L2 writing production, but also due to the variability in the task types investigated, in the operationalisation of task complexity, and in the measures used to explore the effects (Ruiz-Funes, 2015). Also, the lack of sufficient research studies and the inconclusive findings from those that exist have made thematic systematic review of the prior literature on the role of task complexity in L2 production problematic, evidenced in prior studies (Frear & Bitchener, 2015; Ruiz-Funes, 2015) and also in the following sections.

Additionally, as this study is mainly interested in the role of task complexity along the degree of reasoning and the number of elements, in motivation and anxiety, and in pre-task planning in L2 writing production, studies on oral production will not be reviewed. Different processes involved in oral and writing production, different target measures required to be employed in the oral and written modality studies, and the variability in the design of the studies make cross-comparisons and thematic review impractical. As such, in line with some prior studies (e.g., Frear & Bitchener, 2015; Ruiz-Funes, 2015), the available writing studies that can be used in framing this study and discussing the findings are reviewed in the following sections.
2.9.1. Prior Research into Task Complexity along the Resource-directing Dimension

In this study, task complexity along the resource-directing dimension is operationalised by increasing the degree of reasoning and the number of element in monologic writing tasks. As noted above, Robinson (2011a, 2011b) argues that increasing task complexity along the resource-directing dimension in monologic tasks will advantage both syntactic complexity and accuracy, and will have adverse effects on fluency. In this section, previous writing-based studies that have investigated the role of increasing task complexity along the resource-directing dimension are reviewed. First, the studies that have used variables different from the variables of this study and then the studies that have used variables similar to the variables of this study are discussed.

Ishikawa (2007) studied the impact of manipulating task complexity along the resource-directing dimension with regard to the Here-and-Now condition on the narrative writing production of 54 Japanese L2 learners of English. The complexity of the tasks was manipulated by the availability of the cartoon strip and writing in the present versus past tense. For the cognitively simple task, + Here-and-Now, the participants could view the cartoon strip during writing and were asked to write in the present tense, while for the cognitively complex task, - Here-and-Now (There-and-Then), the participants could not view the cartoon strip and were asked to write in the past tense. Ishikawa’s results revealed increasing task complexity along the resource-directing dimension with regard to the Here-and-Now improved the complexity, accuracy, and fluency of L2 learners’ narrative writing production. This study provides a strong support for the CH. However, Ishikawa also allowed five minutes pre-task planning time. It is unknown if the same results would be obtained without
pre-task planning time, as Skehan (2009) has pointed out. In other words, the results might have been conflated due to the simultaneous effects of the resource-directing and resource-dispersing factors.

Kormos (2011) investigated the impact of task complexity on discourse and linguistic features of narrative writing production of upper-intermediate students of English as a foreign language (EFL) in a bilingual high school in Hungary and L1 students of English in the UK. The two narrative tasks provided different levels of cognitive complexity in terms of more/less demand for plot conceptualisation. In cognitively simple task the content was given, but in the cognitively complex task the students had to plan the content themselves. The findings revealed no significant effect for task complexity on linguistic performance except for a major effect on lexical sophistication and a minor effect for temporal cohesion, which was in the opposite direction to the predictions of the CH and did not lend support to LCM either, as the accuracy level remained constant in the performance of the two tasks.

The results of these few studies in L2 writing production, which have operationalised task complexity differently from each other and from this PhD study, are inconsistent. The inconsistency in the findings might be due to the variety in the operationalisation of task complexity and measures employed, and the multiplicity of factors involved, among other things.

In only a few studies is the operationalisation of cognitive task complexity similar to that of my study (see e.g., Kuiken & Vedder, 2007, 2008, 2012; Frear & Bitchener, 2015). However, task types used in these studies are different from what was used in
my study. Letter-writing tasks of varying degrees of complexity were employed in these studies. In my study, argumentative writing tasks were used. It is important to explore the role of task complexity in L2 task performance in different task types, as Skehan’s (2009) review of studies on L2 oral production has revealed the differential impact of task types on L2 production.

In a series of studies, Kuiken and Vedder (2007, 2008, 2011) examined the effect of task complexity of letter-writing tasks operationalised in terms of the number of requirements and types of decisions. The participants’ performance on two letter-writing tasks with different degrees of cognitive complexity displayed a drop in the number of errors (improved accuracy) and a rise in lexical variation on the more complex letter-writing task (Kuiken & Vedder, 2007). Kuiken and Vedder (2008) found a positive effect only for accuracy and no significant effects on syntactic complexity and lexical variation except for French students whose type-token ratio was significantly higher in the complex than in the simple task performance. The effect of task complexity on syntactic complexity for two different proficiency levels, however, was not significant. Kuiken and Vedder (2011) also found a significant positive effect for accuracy and no significant effects on syntactic complexity and lexical variation in the written output.

Recently, Kuiken and Vedder (2012) reported on the results of three studies of task complexity. A significant positive effect was found for accuracy and no effect for syntactic complexity and lexical variation in Study 1. In Study 2, accuracy increased significantly in the complex task performance as a result of the reduction in lexical errors, and contradictory results were obtained for lexical variation: The French
students used less frequent words in the complex task; the reverse was true for the Italian students. The findings of study 3 revealed the effects of task complexity on L2 production was not contingent upon mode (oral versus written production). Similar results were found for oral and written modes. As a function of increasing task complexity, a significant positive effect for accuracy in oral and written modes, no effect on lexical variety in oral and written modes, no effect for syntactic complexity in written mode, and a significant higher syntactic complexity in the simple task performance in oral production were found.

Studies conducted by Kuiken and Vedder (2007, 2008, 2011, 2012) reveal a favourable effect for accuracy, which partially supports the CH. The authors argued that these findings show that the learners’ attentional resources were used for the control of linguistic forms rather than for complexifying linguistic encoding. However, letter-writing tasks were used in these studies and the findings might be partially due to the task type employed. The review of the oral studies shows the differential effects of task types on drawing learners’ attention to different dimensions of L2 production (see Skehan, 2009). Argumentative writing tasks employed in my study may reveal different results.

In a partial replication of Kuiken and Vedder (2007, 2008, 2012), Frear and Bitchener (2015) studied the effect of increasing task complexity on syntactic complexity and lexical variety. Frear and Bitchener added a patently simple task to the two complex letter-writing tasks and measured the ratio of dependent clauses to t-units across all dependent clauses, and the ratio of dependent clauses to t-units with each dependent clause measured separately. Segmental type-token ratio was used to
measure lexical complexity. No significant result was found for the ratio of dependent clauses to t-units across all dependent clauses. When the ratio of dependent clauses to t-units for each dependent clause was analysed separately, significant decreases occurred for adverbial dependent clauses with increases in task complexity: Task 1 (low complexity) had a higher incidence of adverbial clauses than Task 2 (medium complexity) and Task 3 (high complexity).

The result revealed increases in lexical complexity, specifically between Task 1 (low complexity) and Task 3 (high complexity). Apparently, the findings lend support to the CH; however, increases in lexical complexity were associated with no changes or decreases in syntactic complexity, which lends support to the LCM. This indicates that the learners, due to their limited processing capacity, may have prioritised choosing lexical items to formulate the content over the grammatical encoding, similar to what Levelt’s (1989) psycholinguistic speaking model predicts. The authors attributed their insignificant findings to their incorrect alignment of task complexity with the participants’ ability to use automatised subordination rather than viewing the results as disconfirming the predictions of the CH.

In my study, participants were upper-intermediate learners and developmentally sensitive measures of writing syntactic complexity were used. Recruiting upper-intermediate participants might reduce the risk of the incorrect alignment between task complexity and participants’ ability. Additionally, using developmentally sensitive measures of L2 writing syntactic complexity may decrease the possibility of misalignment between the measures and learners’ developmental level. That is, the effect of task complexity can be captured by the dimension of the writing syntactic
complexity that is well aligned with the learners’ developmental level. Previous negative and insignificant effects of task design and implementation features on syntactic complexity might be due to incorrect alignment of syntactic complexity measures and learners’ proficiency level. For instance, using subordination to measure the effect of task complexity for beginner level learners might not be appropriate and may not show favourable effects, as subordination is shown to be the distinctive feature of L2 proficiency in the upper-intermediate level. If significant desirable effects on subordination are obtained for beginner level learners, the results will indicate that these factors stretch learners’ interlanguage system in terms of L2 syntactic complexity rather than only eliciting more instances of previously acquired syntactic forms that are the prominent features of their proficiency level. Additionally, as discussed previously, using developmentally sensitive measures of L2 syntactic complexity can address the potential misalignment issue and indicate whether manipulating task design and implementation features elicits more forms that are the distinctive features of the learners’ current interlanguage system, or whether task design and implementation manipulation also stretches L2 writing syntactic complexity and elicits developmentally more advanced syntactic forms.

In view of the preceding studies in L2 writing and despite some inconsistencies in the findings, increasing cognitive task complexity had a significant positive effect on lexical variation in some studies (e.g., Frear & Bitchener, 2015; Kuiken & Vedder, 2007). For syntactic complexity, Kuiken and Vedder (2007, 2008, 2012) analysed the number of clauses per t-unit and the number of dependent clauses per clause (with dependent clauses analysed as one group), and Frear and Bitchener (2015) examined the ratio of dependent clauses to t-units across all dependent clauses and the ratio of
dependent clauses to t-units for each dependent clause separately. No significant positive effect was found for syntactic complexity in the direction predicted by the CH. A significant positive effect has been found for accuracy in some studies (e.g., Kuiken & Vedder, 2007, 2008, 2011, 2012). Therefore, partial support can be inferred from these results for the predictions of the CH (Robinson, 2001a, 2001b, 2005, 2007, 2011a, 2011b) about the effect of increasing cognitive task complexity on L2 writing production. However, due to the scarcity of research into L2 writing production, further research is warranted to identify the role of task types, task design and implementation features, and the mediating role of individual learner differences in L2 writing production (Robinson 2011a, 2011b, 2015; Skehan, 2009, 2014).

2.9.2. Studies into Task Complexity along the Resource-dispersing Dimension (Planning)

In my study, task complexity along the resource-dispersing dimension is operationalised by providing a 10-minute pre-task planning time in monologic writing tasks. Robinson (2011a, 2011b) argues that reducing task complexity along the resource-dispersing dimension in monologic tasks will advantage syntactic complexity, accuracy, and fluency. In the literature, planning is operationalised as pre-task and online planning (Ellis, 2005). Further, pre-task planning is divided into rehearsal and strategic planning. Rehearsal is performing a task similar to the main task as a preparation to the main task. Strategic planning is further divided into guided and unguided planning. Guided planning refers to focusing on certain feature of task performance, for example, the syntactic structures that can be employed in a task completion, while in unguided planning learners can plan whatever aspect that they intend to (Ellis, 2005). Recently, Skehan (2014) has used preparedness as the umbrella term to conceptualise different types of planning and other factors that
make learners ready to perform the task. In this study, planning is operationalised as an extra 10-minute strategic unguided pre-task planning time. In the following section, previous studies that have investigated the role of planning are reviewed.

Johnson, Mercado, and Acevedo (2012) examined the impact of three distinct forms of pre-task planning (idea generation, organisation, and goal setting) on fluency, grammatical complexity, and lexical complexity of argumentative writing tasks composed by 968 learners of English as a foreign language, a large sample of homogeneous Spanish-speaking EFL learners. Their results revealed that the pre-task planning condition had a small significant impact on writing fluency, but no influence on lexical or grammatical complexity. The authors suggested that the contrasting predictions of the Limited Attentional Capacity Model and the Cognition Hypothesis may be limited to oral production and may not be applicable to writing production. They also argued that the findings in pre-task planning in previous L1 and L2 studies might have been affected by the participants’ education and genre knowledge, and that for pre-task planning to show its impact on L2 writing production a threshold level of general L2 proficiency may be required.

Kroll (1990) examined the impact of writing under two conditions (within a 60-minute time limit in class and writing without a time limit at home) on L2 writers’ composition addressing the question “What does time buy?” Her results showed that when L2 writers’ wrote at home their accuracy in composition improved, and they received a higher rating on their composition. The difference between compositions produced at home and in the class was not found to be significant. However, as Kroll herself acknowledged, the planning behaviours for the two conditions and the amount
of time L2 writers spend on planning and composing while writing at home were not controlled in her study.

Friedlander (1990) studied whether planning in the language in which L2 writers had acquired the information about the topic improved their planning and the content of the writing compositions. The impact was positive. However, the overall ratings of the planned essays were not statistically significant regardless of the language writers used for planning. Using think-aloud protocols, Whalen and Menard (1995) studied L1 and L2 writers’ planning in their L1 and L2. They found no significant difference in their planning quality in the two languages. However, they found that linguistic and pragmatic planning occurred more often in the writers’ L1 writing. De Larios, Marin, and Murphy (2001) explored within-task planning. They found that formulation (i.e., translation in Kellog’s 1996 model) played a leading role in the argumentative writing production in relation to other composing processes such as planning regardless of writing in L1 or L2. Higher proficiency L2 writers, however, allocated less time to formulation and were able to devote more time to planning. These studies did not report the impact of planning on CALF.

Ellis and Yuan (2004) argued that planning in L1 or L2 is not important unless the topical knowledge is acquired in L1, that L2 leaners’ attention to translation may deprive them of on-line conceptual content planning, and that the effect of planning on writing production is unknown due to the lack of clear distinction between pre-task and on-line planning and use of holistic ratings rather than specific measures of writing production (CAF) in prior research. Therefore, Ellis and Yuan explored the differential impact of pre-task, unpressured on-line planning, and no planning on the
CAF of 42 Chinese learners’ narrative writing production. Their results showed that pre-task planning assisted greater fluency as measured by syllables per minute of overall writing time and greater syntactic variety as measured by the number of different verb forms, whereas the unpressured on-line planning condition resulted in greater accuracy as measured by error-free clauses. The authors concluded that three conditions of planning affected different aspects of writing processes differently. They proposed that the no-planning condition had negative consequences for L2 writing production as the L2 writers had to attend to formulation, execution, and monitoring in writing under pressure. The pre-task planning provided better opportunities for formulation and unpressured on-line planning promoted monitoring which resulted in positive consequences for fluency and syntactic variety, and accuracy for the two planning conditions, respectively.

The findings of Johnson, Mercado, and Acevedo (2012) differ from those of Ellis and Yuan (2004) due to huge differences in the design of the two studies. While Ellis and Yuan’s (2004) findings are promising, these two studies are unable to ascertain the impact of task design features and individual differences from the research design employed. Therefore, my study will extend these studies by investigating the simultaneous impact of task design features and pre-task planning conditions on L2 writing production. However, as the purpose of my study is to compare the contrasting predictions of the TOH and the CH, I will exclude the unpressured online-task planning from the research design.

Ong and Zhang (2010) examined the impact of manipulating task complexity on argumentative writing production of 108 Chinese L2 learners of English. Tasks were
manipulated with regard to the availability of pre-task planning time, the provision of ideas and macrostructure of the writing tasks, and the availability of the draft during revision. Ong and Zhang’s (2010) results revealed: (a) increasing task complexity along the planning time continuum led to significantly greater fluency II (i.e., mean number of words produced per minute of the total time spent on the task) and lexical complexity; (b) increasing task complexity by not providing ideas and macrostructure led to significantly greater lexical complexity but had no effect on fluency I (i.e., mean number of words produced per minute of transcription) or fluency II, and (c) increasing task complexity by draft unavailability did not lead to significant differences in fluency and lexical complexity. Ong and Zhang (2010), however, examined the impact of task complexity on lexical complexity and fluency. The trade-off effect and the impact of task complexity on other aspects of writing production were not reported. The findings with regard to the unavailability of pre-task planning time are in opposition to the predictions of the CH with regard to the impact of the resource-dispersing dimension on L2 production. However, the small simple size of this study and the lexical complexity measure, the type–token ratio, which is shown to be sensitive to produced sample (written text) length (MacWhinney, 2002), might have influenced the findings.

Ong and Zhang (2013) posited that little is understood about the differential impact of different planning and revising conditions on text quality produced by L2 writers. They explored the impacts of various planning conditions, sub-planning conditions, and revising conditions on 108 Chinese English-as-a-foreign language (EFL) learners’ argumentative writing text quality, as measured by Jacobs, et al.’s (1981) ratings scale rather than specific measures of writing production. They found that
free-writing condition assisted FL learners’ written text quality, while both the task-content given condition and the task-content-organisation-given condition resulted in significantly better writing quality than the task-given condition.

With regard to the initial-essay-accessible and initial-essay-removed conditions, there was no significant difference in the produced text quality. They concluded that the free writing condition improved the overall text quality by facilitating content retrieval, and that the cognitive load imposed on the FL writers’ working memory resources by the writing task was effectively reduced in the task-content-given and the task content-organisation-given conditions. The main concern of this study was written text quality. It did not examine the impact of task complexity on writing production (CLAF), and does not contribute to resolving the opposing predictions of the two psycholinguistics models of task complexity and linguistic production.

In view of the above section regarding the role of pre-task planning in L2 writing production, positive effects on fluency and complexity can be gleaned from the findings of the studies that have operationalised planning as an extra 10 minutes of pre-task planning and measured the effects on CAF (e.g., Ellis & Yuan, 2004; Johnson, Mercado, & Acevedo, 2012). These findings partially support the predictions of the CH and the LCM. Nonetheless, the review of the limited research on the impact of planning on L2 writing production suggests that the interface of task complexity and L2 writing production is still in its embryonic stage. Also, because of the multiplicity of factors involved, the comparison of the findings across the studies is difficult and the cross-case comparison of such studies may well be misleading. Additionally, the simultaneous effects of task complexity along the resources-
directing dimension and planning have not been explored in L2 writing extensively (but see Frear, 2013). As such, we do need further research to explore the isolated, simultaneous, and mediating role of the various aspects of task complexity, including task types and learner individual differences in L2 writing production. Also, we need to conduct full and partial replications of the prior research studies to make comparisons and, in turn, theorising possible. Therefore, concluding that theories of task complexity are not applicable to writing production, based on such a limited number of studies that have investigated different aspects of task complexity with totally different task types and different levels of proficiency without examining the mediating role of individual learner factors, may be premature (Johnson, Mercado, & Acevedo, 2012).

2.9.3. Studies into Task Complexity and Second Language Writing Assessment

The effect of different task types on test scores have also been studied in the field of second language writing assessment. Schoonen (2005) found that the type of text students produced in a writing test had a greater effect on scores than the learners’ writing ability, which highlights the importance of examining particular features of the task that might contribute to the variance in L2 writing performance. In another study, Hamp-Lyons and Mathias (1994) found that students achieved higher scores in tasks that were judged as more difficult by expert raters than in tasks that were deemed easier. They argued that cognitively complex writing tasks more likely motivated learners to perform better than did cognitively less complex tasks. Further research is warranted to identify the effects of task type, and task design and implementation features on learners’ performance to provide potential insights into task-based writing assessment, as the potential effects of these factors are evidenced in the findings of these two studies.
2.9.4. Studies into Simultaneous Effect of Task Complexity along the Resource-directing and the Resource-dispersing Dimensions

There is only one study on the simultaneous effect of task complexity along the resource-directing and the resource-dispersing dimensions in terms of L2 oral production (Gilabert, 2007), and there is one relating to L2 writing production (Frear, 2013). I will review both studies here due to scarcity of research into the simultaneous effects of task complexity and planning conditions on L2 writing production. In fact, the scarcity of research along this line of inquiry further contributes to the significance of this PhD study that examined both the isolated and simultaneous effects of task complexity on L2 writing production.

Gilabert (2007) examined the effects of pre-task planning conditions and the degrees of displacement on CAF of 48 first and second-year L2 learners’ oral production. He compared the L2 oral production with regard to accuracy, fluency and complexity under four conditions: Pre-task planning Here-and-Now, pre-task planning There-and-Then, and No-pre-task planning Here-and-Now, and No-pre-task planning There-and-Then. Gilabert found that simple Here-and-Now and complex There-and-Then tasks performed under a 10-minute planning condition significantly enhanced fluency. Regarding the effect of four planning conditions on syntactic complexity and accuracy, no significant differences were found. In all performance conditions, the planned conditions triggered a significantly higher lexical diversity than the unplanned ones. Therefore, the planning conditions resulted in better fluency and lexical diversity, but not in better accuracy or syntactic complexity.
Regarding Here-and-Now and There-and-Then conditions, learners performed narrative tasks significantly more fluently in Here-and-Now than There-and-Then conditions under both the planned and unplanned conditions. As far as lexical and syntactic complexity is concerned, contrary to the predictions of Robinson’s Hypothesis, increasing task complexity (in There-and-Then condition) reduced learners’ performance indexes. However, learners’ accuracy improved significantly in There-and-Then condition (cognitively complex condition) under both the planned and unplanned conditions. Gilabert (2007) concluded that simultaneous focus on accuracy and complexity is achievable if tasks are manipulated to be cognitively complex along the resource-directing dimension, but simple along the resource-dispersing ones. He further argued that task complexity is a strong and testable construct, and that the findings of task complexity studies may have pedagogical applications for task-based as well as process and content-based approaches to the syllabus design and teaching second languages.

In L2 writing, Frear (2013) studied the simultaneous effects of increasing task complexity along the resource-directing dimension and 10 minutes of pre-task planning. For syntactic complexity, providing pre-task planning in the complex writing task rendered no significant result when all subordinate clauses were considered as one group. However, when the effect on adjectival, nominal, and adverbial clauses was analysed separately, a significant positive increase was found for the adjectival dependent clauses in the complex task performance. Nonetheless, the significant increase in syntactic complexity coincided with nonsignificant changes in lexical complexity that partially supports the CH. In these two studies, Gilabert (2007) examined Here and Now versus There and Then, and Frear (2013)
used letter-writing tasks and examined the effects on complexity with regard to increases in the use of subordination and on lexical complexity in terms of lexical diversity as measured by segmental type-token ratio. In Frear’s study, the effects on other dimensions were not studied. The authors acknowledged the possibility of misalignment problem between learners’ proficiency level and the complexity of the tasks, and subordination was defined as the advanced form of syntactic complexity in L2 writing. As discussed above, the effects of task complexity on L2 writing production in argumentative writing tasks might be different. In L2 writing, sub-clausal complexity, not subordination, is considered as the advanced syntactic form. The effects on accuracy, fluency, lexical sophistication, content, organization, and writing quality are also needed to be investigated in order to provide an in-depth documentation of the phenomenon.

2.10. Writing Motivational Beliefs, Writing Anxiety, and Writing Task Complexity

Prior research into the effect of task complexity on L2 writing production has not examined the mediating effect of individual learner differences (ID). In fact, prior research (e.g., Ellis & Yuan 2004, Kormos, 2011; Kuiken & Vedder, 2007, 2008, 2011, 2012; Ong, 2013, 2014; Ong & Zhang, 2010, 2013) calls for the inclusion of ID factors in future research. Robinson (2011a, 2011b) has hypothesised that the role of affective factors will be more clearly manifest in performance on the cognitively complex task. To occupy this gap in prior research, I examined the extent to which writing motivation and anxiety modulate the effect of task complexity in L2 writing production. In this study, motivation and anxiety are operationalised as multidimensional constructs that further contribute to the significance of the study.
Operational definitions and prior research into writing motivational beliefs and anxiety are discussed in the following sections.

2.10.1. Motivational Beliefs and Writing Task Complexity

The cognitively challenging nature of writing is widely acknowledged by scholars. Being inherently an arduous and cognitively demanding task, writing requires a high level of motivation. In fact, writing models (e.g., Hayes, 2012; Hayes & Nash 1996; Zimmerman & Risemberg, 1997) acknowledge the significant role of motivation in writing development and performance. In second language acquisition research, the role of motivation is widely acknowledged, and Dörnyei (2009) has proposed task motivation comprises general (trait) and specific (state) motivation that influence task commitment. The former relates to the persons’ overall predisposition that are believed to be constant and enduring, while the latter refers to the learners’ temporary reaction to the task.

Motivation is a dynamic system and a trait/state conceptualisation of motivation does not account for the dynamic nature of motivation (Dörnyei; 2002, 2009). Therefore, Dörnyei (2009) has developed the motivational task processing system that identifies three motivational phases: Pre-action phase where motivation is initiated and triggers the starting of the task, action phase where the motivation needs to be sustained and results in commitment to task performance, and post-action phase that involves reflection on the performance that might influence future task involvement and commitment. Although the current study acknowledges the dynamic nature of motivation, it focuses on the learners’ general writing predisposition drawing on L1 writing motivation research. Learners’ general writing motivation has been shown an
important factor in writing task performance in L1 writing research (Troia, Shankland, & Wolbers, 2012).

In writing research, motivation is acknowledged as a necessary requirement for successful writing development and performance (Pintrich & Schunk, 2002). However, it is acknowledged that motivation is a multidimensional rather than a unitary concept that comprises self-efficacy beliefs, goal orientations, success attributions, perceived task value, interest, and attitudes (Troia, et al., 2012). Hence, a multidimensional conceptualisation of motivation is adopted in this study to investigate the strength of the correlations between the dimensions of motivation and L2 writing production in simple versus complex versions of the task. The review of research into the effect of motivational beliefs components on writing development and performance is presented in the following sections.

2.10.1.1. Writing Self-efficacy Beliefs

Bandura’ (1997) conceptualisation of self-efficacy, which refers to individuals’ assessment of their ability to complete a task, is likely the most well researched and well established facet of human motivation (Troia, et al., 2012). An individual’s assessment of his/her competence to perform a task, self-efficacy belief, is largely shown to associate with the human adaptive behaviour irrespective of age, sex, or ethnic variability (e.g., Bandura, 1997; Pajares, 1996; Pintrich & DeGroot, 1990; Pintrich & Schunk, 2002). The documented adaptive behaviours comprise the extent of effort expended completing a task, perseverance in case of difficulties and challenges, and strategies used to perform a task successfully. Bandura (1997) and
Eccles and Wigfield (2002) conceptualise that self-efficacy belief consists of two constructs: Outcome expectations and efficacy expectations.

Outcome expectations refer to an individual’s assessment that his/her particular actions will result in desired outcomes and goals, and efficacy expectations refer to individuals’ assessment of their competence in performing the requisite action to accomplish the desired outcomes and goals. For instance, some students might believe that they can successfully meet the requirements for the successful completion of a PhD programme, but not necessarily that their successful completion of the PhD programme will lead to their desired outcome — a career in academia.

As regards the role of self-efficacy in writing, prior research has established that self-efficacy plays a prominent role in writing performance (see Pajares & Johnson, 1994, 1996; Pajares, Miller, & Johnson, 1999; Pajares & Valiante, 1997, 1999; Shell, Colvin, & Bruning, 1995; Shell, Murphy, & Bruning, 1989).

2.10.1.2. Writing Task Value

According to expectancy-value theory (see Hidi, Berndorff, & Ainley, 2002), task interest or value is another primary construct of motivation, which reflects the task importance or task value for the individuals (Schiefele, 1999; Wigfield, Eccles, &1992). Task value affects the goals individuals set and the adaptive behaviours they embrace including more attention, effort, engagement, enjoyment, persistence, and learning (e.g., Schiefele, 1991). Not surprisingly, research has demonstrated that task value, in conjunction with self-efficacy belief, plays a facilitative role in writing performance, (Albin, Benton, & Khramtsova, 1996; Benton, Corkill, Sharp, Downey, & Khramtsova, 1995, among others). Given that prior research has demonstrated
high self-efficacy belief and high task value result in more attention, more effort, and perseverance in performing difficult tasks, in line with Robinson (2011a, 2011b), I hypothesise that the mediating role of self-efficacy belief and task value will be more clearly manifest in the cognitively complex task than in the simple one.

2.10.1.3. Writing Success Attribution

The other component of motivation is attribution, which refers to the causes to which individuals attribute their successes and failures (Weiner, 1986). Individuals’ perception of causes of their success and failure is affected by the perception of the amount of their control over the cause, its locus, and its stability (Weiner, 1986). Individuals who attribute their success and/or failure to factors under their own control, such as sustained sufficient strategic effort, and failures to either insufficient strategic effort or to unrealistic goals (Weiner, 1986) are more likely to embrace adaptive motivational beliefs and behaviours. That is, these individuals, regardless of whether they fail or succeed, believe that by expending sufficient strategic effort and setting realistic achievable goals they can improve their own performance outcomes.

Conversely, those who attribute their performance outcomes to external uncontrollable factors such as being lucky or unlucky, task ease or difficulty, or being intelligent or having limited ability are more likely to form maladaptive motivational and behavioural patterns (Schunk, 1984). Although research on the role of attribution in writing performance is scant (Troia, et al., 2012), with regard to the effect of interaction between attribution and task complexity, I hypothesise that the modulating effect of effort attribution is more likely to be more clearly reflected in the cognitively complex task performance, as individuals with maladaptive
motivational beliefs are less likely to be motivated to perform well on cognitively complex tasks. These individuals believe that their efforts have little or no influence on their performance outcomes, as they typically attribute their performance outcomes to external factors such as task difficulty.

2.10.1.4. Writing Goal Orientation

Goal orientation, another component of human motivation, is divided into mastery, performance approach (Ames, 1992; Middleton & Midgley, 1997), and performance avoidance goals (e.g., Senko, Hulleman, & Harackiewicz, 2011). Individuals pursuing mastery goals are inclined to attain knowledge and skills and a sense of competence and learning, whereas those pursuing performance approach goals are interested in public recognition of their skills and abilities, performing better than others and receiving positive evaluation (Ames, 1992). Conversely, individuals with avoidance approach goals might avoid performing tasks in order not to exhibit their perceived incompetency.

Mastery goals are shown to be associated with many positive learning attributes, including higher self-efficacy, greater self-regulation, and better achievement (see Ames, 1992; Pintrich & DeGroot, 1990). A performance approach goal is not considered to be necessarily maladaptive (e.g., Pajares, Britner, & Valiante, 2000), although it is unknown for which learners and under what conditions performance goal is not maladaptive (Midgley, Kaplan, & Middleton, 2001). Given that learners pursuing mastery goals embrace challenging tasks, writers with mastery goals will more likely perform better on complex writing tasks compared to those pursuing performance approach or performance avoidance goals.
2.10.2. Anxiety and Writing Task Complexity

Writing anxiety, an individual learner factor, defined as “fear of the writing process that outweighs the projected gain from the ability to write” (Thompson, 1980, p. 121, quoted in Jahin, 2012) is shown to affect writing performance negatively (Atay & Kurt, 2006; Pajares & Johnson, 1994, among others) and also the writing process (Cheng, 2004; Huang, 2009; Lee, 2005, among others). However, one should note that earlier research on the relationship between anxiety and language performance resulted in contradictory findings (e.g., Chastain, 1975) which made their comparisons difficult (Cheng, 2004; Gardner & MacIntyre, 1993; Scovel, 1978; Zhang, 2000, 2001). Many researchers have attributed these inconsistent results in part to inadequate conceptualisations, operationalisations, and measures of L2 anxiety (e.g., Cheng, 2004; Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1989; Zhang & Rahimi, 2014).

To address the discrepant results, obtained during the mid-1980s and later, L2 anxiety was conceptualised as a unique, situation-specific form of anxiety, which emerges in response to L2 learning rather than as an indication of the more general type of anxiety (Horwitz et al., 1986; Gardner, 1985; MacIntyre & Gardner, 1991a). Further, Gardner and MacIntyre (1993, p. 5) defined L2 anxiety as a stable personality trait “as the apprehension experienced when a situation requires the use of a second language with which the individual is not fully proficient”.

With the conceptualisation of L2 anxiety as a specific anxiety and subsequent developments in L2 anxiety measures, Gardner and his colleagues’ French Class Anxiety Scale and French Use Anxiety Scale (MacIntyre & Gardner, 1989), and
Horwitz et al.’s (1986) Foreign Language Classroom Anxiety Scale were developed (see also Sparks & Patton, 2013). Many correlation studies have shown the negative impact of anxiety in L2 performance (e.g., Aida, 1994; Cheng, Horwitz, & Schallert, 1999; Gardner, Lalonde, Moorcraft, & Evers, 1987; Phillips, 1992; Young, 1986). Moreover, many studies in a controlled environment have confirmed the negative impact of L2 anxiety on L2 learning processes (MacIntyre & Gardner, 1991b, 1994a, 1994b).

Subsequent developments in L2 anxiety research and skill specific conceptualisation and measures of anxiety have shown a consistent relationship between FL and writing apprehension and various measures of FL and SL writing processes, performance, and quality (e.g., Atay & Kurt, 2006; Cheng, Horwitz, & Shallert, 1999). As writing is a solitary journey and teachers primarily provide feedback on the product of learners’ writing, depending on the nature of feedback, which might be vague and general and may not provide specific strategies for improving writing, writers might feel confused, passive, or frustrated (Williams, 2003). Writers might also feel that they do not have access to support and encouragement (Tsui, 1996).

Although because of developments in the conceptualisations and measures of L2 anxiety, research has resulted in consistent findings on the negative relationship between L2 anxiety and L2 performance, Cheng (2004) argued that previous measurers of writing anxiety need further development. He developed and validated a three-dimensional Second Language Writing Anxiety Inventory (SLWAI) comprising three subscales: Somatic Anxiety, Cognitive Anxiety, and Avoidance Behaviour. Cheng’ reliability coefficients, correlation, and factor analysis revealed
that the total scale and subscales of the SLWAI have acceptable reliability and validity. As such, I employed this inventory to further illuminate the role of writing anxiety in L2 writing performance.

Anxiety is not a unitary construct, but a multidimensional phenomenon that consists of different response aspects. In fact, Cheng (2004) has convincingly argued that accumulating research on both antecedents of anxiety and the impact of anxiety on diverse facets of human behaviour or academic performance (Morris, Davis & Hutchings, 1981; Smith & Smoll, 1990) has confirmed the multidimensional conceptualisations of anxiety. In fact, several multidimensional measures of anxiety have been developed: Test anxiety (Morris et al., 1981; Sarason & Sarason, 1990), speech anxiety (Fremouw & Breitenstein, 1990), and sport performance anxiety (Smith & Smoll, 1990). In these measures, anxiety is conceptualised, operationalised and classified into several independent dimensions, encompassing somatic or physiological (e.g., upset stomach, pounding heart, and excessive sweating,), cognitive (e.g., worry, preoccupation, and negative expectations), and behavioural (e.g., procrastination, withdrawal, and avoidance) aspects.

Prior research following a multidimensional conceptualisation and measurement of anxiety has demonstrated the differential effect of the components of anxiety on performance (see Bishop, Holm, Borowiak, & Wilson, 2001; Deffenbacher, 1977). Obviously, this significant finding would have not been revealed by a one-dimensional conceptualisation and measure of anxiety. Therefore, I used Cheng’ (2004) SLWAI to investigate the interaction effect of three components of anxiety and task complexity on FL writers’ text.
Steinberg and Horwitz (1986) posit that final grades are not sufficiently sensitive to anxiety and recommend that a range of outcome measures be used in studying the effect of L2 anxiety on L2 production. Indeed, the results of their own study indicated that L2 anxiety affected the content of the L2 speech production. As such, motivated by prior research on L2 anxiety and task complexity research, multiple measures of L2 writers’ writing production will be employed to address the interface between task complexity, L2 writing anxiety, and L2 writing production. Following a multidimensional conceptualisation and measurement of anxiety, I am interested in understanding whether components of anxiety have facilitative or debilitative modulating roles in the simple versus complex writing task performance as measured by CALF, content, organisation, and written text quality.

2.11. Summary of the Literature Review

As discussed in the Introduction chapter and in this chapter, TBLT has gained popularity due to its theoretical strength and research evidence (e.g., Ellis, 2013; Skehan, 2014; Long, 2015; Robinson, 2011a, among others), and tasks have gained a central role in SLA research, second language teaching, and L2 writing pedagogy (Hyland, 2003). However, as Long posits, there are unresolved issues in TBLT. Frameworks have been proposed to address the remaining issues, and studies have been conducted to provide insights into developing principles for task-based syllabus design, methodology, and assessment. Nonetheless, as the Literature Review revealed, the results from a few existing research studies into the effect of task complexity on L2 writing production are inconclusive, and the mediating role of affective factors in L2 writing task performance has not been explored to date. Also,
research into the synergistic impacts of increasing task complexity along the level of reasoning and number of elements, and planning conditions is scarce. This study endeavours to contribute to filling these lacunae and building research-based principles for task-based L2 pedagogy by exploring task design and implementation features mainly by examining Robinson’s (2011a, 2011b) hypothesis. Some dimensions of Robinson’s TCF for task complexity, his rationale for sequencing tasks from simple to complex, and some predictions of the CH are examined in this study.

Specifically, this study investigates six claims of the CH, namely: (a) Increasing the level of reasoning and number of elements along the resource-directing dimension will increase cognitive task complexity; (b) increasing task complexity along the level of reasoning and the number of elements will result in the enhancement of syntactic complexity and accuracy, and will have adverse consequences for fluency; (c) providing pre-task planning will advantage CAF; (d) increasing cognitive complexity of tasks will increase using developmentally more advanced forms of language; (e) the mediating role of affective variables will be more clearly manifest in cognitively complex task performance, and (f) increasing task complexity along the level of reasoning and the number of elements, and providing pre-task planning will have positive synergistic effects on CAF.
Chapter 3.

Research Methodology and Design

3.1. Chapter Overview

In this chapter the research objectives, questions, hypotheses, and research design are presented. After describing the research foci, the pilot study is described; this is followed by the description of the participants, context of the study, research design, data collection procedures, and materials used to collect the data. The independent and dependent variables of the study and the measures used for data coding and rating are described next. The chapter concludes with presenting the reliability coefficients of data coding and rating.

3.2. Methodology

Methodologically, this research study was undertaken within a pragmatic paradigm that employs pluralistic strategies of research inquiry depending on the knowledge type required to address the research questions (Creswell, 2009). To elaborate, the study is mainly grounded in a positivist/normative research paradigm (Ellis & Barkhuizen, 2005). However, an interpretative paradigm was also employed to evaluate the validity of the two designed tasks. To this end, iterative readings (Patton, 2002; Dörnyei, 2007) of the teachers and the learners’ reasons for their judgements of the difficulty of the two writing tasks, guided by the research question, were performed.
The positivist/normative research paradigm (Ellis & Barkhuizen, 2005) typically adopts quantitative strategies, comprising experimental and associative categories (Mackey & Gass, 2005). The experimental strategy is employed to establish causality between the independent and dependent variables. The associative quantitative strategy does not seek to establish causality between the variables of the study; it is used to determine whether there is a correlation between the variables of the interest, and if so, the direction and the strength of the correlation. The experimental strategy was used to explore the effect of task complexity and pre-task planning on the target measures of L2 writing production in Study 1 and 2. The quantitative associative strategy was utilised to research the mediating role of motivation and anxiety in the simple versus complex writing task production in Study 1. Additionally, open-ended questions were used to collect qualitative data and learners’ self-rating to collect questionnaire data in order to assess the validity of the two employed tasks, operationalised as simple versus complex.

3.3. Scope of the Research

Although tasks have a central role in L2 learning and L2 writing pedagogy (Hyland, 2003), establishing criteria for task complexity and subsequently designing and sequencing pedagogical tasks to form comprehensive L2 programmes are unresolved issues in TBLT (Long, 2015, 2016). Additionally, research into writing task design and implementation features and learner ability and affective factors that can provide potential insights into pedagogical task design and sequencing and, ultimately, developing comprehensive L2 programmes is scarce, but emerging (Ellis & Yuan 2004; Frear & Bitchener, 2015; Kormos, 2011; Kuiken & Vedder, 2007, 2008, 2011, 2012; Ong & Zhang, 2010, 2013; Ruiz-Funes, 2015). However, the results are far
from conclusive mainly due to the variety in research designs of prior studies and the multiplicity of factors involved.

To sum up, research on task complexity and writing production is at an embryonic stage and due to the multiplicity of factors contributing to the variations in task-based writing production, drawing conclusions and making generalisations require an extensive body of research from multiple perspectives that is not accessible to date. To contribute to nascent body of research, the current thesis documents research into the effects of task complexity, task conditions, and task difficulty on complexity, accuracy, lexical diversity, and fluency (CALF), and content, organization, and writing quality in FL/L2 writing. The specific research objectives and questions are presented in the following sections.

3.4. Research Objectives

The following research objectives were set in the light of inconclusive findings from previous research about the impact of task complexity on L2 writing production; of several gaps in research about the effect of task complexity on writing content and organisation; about the simultaneous impact of the resource-directing and the resource-dispersing dimensions on L2 writing, and about the moderating impact of individual difference (ID) factors on L2 writing production. The specific objectives are as follows:

(a) to investigate the isolated impact of task complexity on writing production (i.e., CALF, content, organisation, and text quality) of upper-intermediate level foreign language (FL) writers of English (Study 1),
(b) to explore the impact of task complexity on L2 learners’ perception of task difficulty (Study 1),
(c) to explore the extent to which L2 learners’ writing production (i.e., CALF, content, organisation, and text quality) is linked to the level of the learners’ motivation as indicated by their self-report on a Motivational Beliefs questionnaire (Study 1),
(d) to explore to the extent to which L2 learners’ writing production (i.e., CALF, content, organisation, and text quality) is linked to the level of the learners’ writing anxiety as indicated by their self-report on Second Language Writing Inventory (Study 1), and
(e) to study the simultaneous effect of task complexity and task planning condition (present Vs. absent) on writing production (i.e., CALF, content, organisation, and text quality) of upper-intermediate FL writers of English (Study 2).

3.5. Research Questions

This study addresses the following research questions:

1. What is the effect of task complexity (less vs. more) on L2 writing production (i.e., CALF, content, organisation, and text quality)?
2. What is the effect of task complexity (more vs. less) on perceived task difficulty?
3. What is the relationship between each dimension of learners’ writing motivational beliefs’ and each measure of L2 writing production (i.e., CALF, content, organisation, and text quality)?
4. What is the relationship between each dimension of learners’ anxiety and each measure of L2 writing production (i.e., CALF, content, organisation, and text quality)?

5. What is the simultaneous effect of task complexity and planning (present vs. absent) conditions on L2 writing production (i.e., CALF, content, organisation, and text quality)?

3.6. Research Hypotheses

Concerning the effect of task difficulty on L2 production (research question 1), Skehan and Foster (1999, 2001) anticipated that learners perform better on less-difficult tasks, as they do not have to allocate their entire limited attentional capacity to the content of the task, and thus they can devote part of their attention to linguistic forms. Skehan and Foster postulate that as learners’ attentional resources are limited, there will be a trade-off between content and form, as the difficulty of the task increases. Contrary to this hypothesis, Robinson (2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b) proposes that learners perform better on complex tasks, as learners are able to share their multiple attentional resources between content and form. Robinson predicts that increasing the cognitive complexity of tasks results in more awareness and leads to greater syntactic complexity, more lexical variation, and higher accuracy on complex task performance.

With regard to the simultaneous effects of task complexity along resource-directing dimension and pre-task planning, Robinson’s (2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b) Cognition Hypothesis postulates that planning will allow learners to better utilise their multiple attentional resources and perform better on the more
complex task in terms of syntactic complexity, lexical variation, and accuracy. However, in a no planning condition, Robinson proposes, learners’ performance will deteriorate, as they will not have sufficient time to use their multiple attentional resources. In terms of the moderating effects of motivational beliefs and anxiety on the learners’ performance, the Cognition Hypothesis predicts that the role of these factors will be more clearly manifest in the complex task performance than in its simple counterpart. Findings by Révész (2011) on the effect of task complexity on oral production disconfirmed this hypothesis, as there was no significant correlation between anxiety and L2 oral production in her study. To date, however, the moderating effect of multidimensional conceptualisations of motivational beliefs and anxiety on L2 writers’ task performance on the complex versus simple versions of tasks has not been investigated. In view of the gaps in research on task complexity and L2 writing production and discrepancies in the results of a few extant studies, this study examines the predictive power of both Robinson’s Cognition Hypothesis and Skehan’s Trade-Off Hypothesis in L2 writing production.

3.7. Pilot Study

The pilot study was conducted for six main reasons: First, to ascertain that the writing tasks elicit writing samples of sufficient length from each participant for coding and rating purposes; second, to obtain 10 teachers’ written judgments of the complexity of the two tasks for their students and their reasons for their judgments; third, to verify learners’ perceptions of the complexity of the adapted tasks in relation to the researcher’s operationalisation. It is acknowledged that there exists prior research which has already provided sufficient evidence for this purpose in the fields of SLA (e.g., Prabhu, 1987; Révész, 2011; Robinson, 2001a, 2005) and in cognitive
psychology (e.g., Halford, Cowan, & Andrews, 2007); fourth, to identify the mean
time spent on completing the writing tasks; fifth, to examine the clarity and
comprehensibility of the questionnaires and writing tasks instructions and wordings
to the learners and to make any required amendments prior to conducting the main
study and, finally, to collect sufficient authentic written data for the researcher and
research assistant to practice coding authentic data and to rate learners’ writing on a
100-point scale for the purpose of achieving satisfactory levels of intra-rater and
inter-rater reliability. The same procedure was used for conducting the pilot and the
main study.

3.7.1. Piloting the Questionnaires

To ensure the clarity of the instructions and items of the questionnaires to the
participants and to identify the maximum amount of time needed to complete the
questionnaires, the questionnaires were piloted with 10 upper-intermediate EFL
learners from the same context, students learning English in private language schools
in central Iran. The participants were instructed to raise any questions and difficulties
they faced in completing the questionnaires. The instructions and the items of the
questionnaires were clear to the participants except that some of the participants had
difficulty understanding the meaning of some words. Based on this finding,
amendments were made to the wordings of the questionnaire items by inserting a
synonym in brackets for those words identified as difficult. Also a decision was made
to allow the participants to use their bilingual dictionaries and to ask any clarifying
questions during the completion of the questionnaires in the main study. The
questionnaires were not translated into Farsi. Upper-intermediate L2 learners could
find Farsi version of the questionnaires demotivating, and translation of the questionnaires could change the intended meaning of the items.

As regards the time limit, the slowest student completed the general information questionnaire and individual difference questionnaire in 35 minutes, including the time allowed for checking the meaning of the new words. The longest time spend on completing each task difficulty questionnaire was 5 minutes. Accordingly, these time limits were set for the completion of the questionnaires in the main study.

3.7.2. Piloting the Writing Tasks

To establish the difference in the difficulty level of the two writing tasks, to verify the clarity of the instructions, and to determine time needed to complete the two writing tasks, the two writing tasks were piloted with the same 10 participants who participated in piloting the questionnaires. Also 10 teachers, who were teaching in the private language schools, where this study was conducted, were recruited on a voluntary basis to judge the difficulty level of the two writing tasks for their upper-intermediate level students.

All the students and teachers in the pilot study and the vast majority of participants in the main study ($n = 129$) 92% found the writing task which required allocating 10,000,000 million dollars for six competing projects more difficult and challenging than the simple task which instructed the participants to allocate 5,000,000 million dollars for three competing projects. This finding verified the researcher’s operationalisation of the two writing tasks. The majority of participating students and all teachers believed that prioritising six competing projects and allocating a larger
amount of total fund were more difficult than prioritising and allocating funds for three competing projects. They believed that the projects were all at a similar level of worthiness for funding and worthy of attention, and allocating funds for one affected the others: what is termed elements trade-off. There were more projects and more elements trade-off in the six-project task. That is, increasing the number of elements increases task complexity, and if there is a trade-off between the elements, the task becomes even more complex. This finding was corroborated by the participants’ reasons for the complexity of the six-project task compared to that of the three-project task in the main study (see Chapter 4 for the findings of the main study).

The instructions of the writing tasks were clear to the participants, as they were allowed to consult their bilingual dictionary to assist their comprehension. This decision was made to avoid confounding the findings due to potential differences in the participants’ reading comprehension ability. Initially, the participants were instructed to prioritise and allocate the funds without emphasising that they should allocate funds to all projects. Some participants allocated funds to only two projects in each task, so an amendment was made which instructed the participants to prioritise and allocate funds to all competing projects.

As regards time allocation for each task, initially it was decided that each participant should spent 30 minutes for each task, as this was the time the participant were typically given to complete their writing tasks in their classes. However, piloting results indicated that the slowest participant spent three minutes to comprehend each writing instruction and 32 and 35 minutes to complete the simple and complex writing tasks, respectively. As the successful completion of the tasks by all
participants was the main concern of the researcher (as in Yuan & Ellis, 2003), a decision was made to allocate three minutes for reading and comprehending each writing instruction and 35 minutes for writing each writing task.

3.7.3. Piloting Data Coding and Scoring

Fortunately, the writing tasks elicited sufficient text length from the participants; some participants wrote even lengthier texts than required. The texts were coded and scored by two raters well-versed in coding and scoring L2 writing production by using measures of CALF and Jacob et al.’s (1981) analytic writing scheme. The two raters coded and scored two texts together to reach an agreement on the criteria for measures of complexity, accuracy, and analytic scheme of writing production. Then, the two raters coded and scored eight remaining texts separately. Finally, the raters compared their coding and scoring which resulted in their clear grasp of the coding and scoring procedures and criteria. The Spearman’s rho correlation for inter-rater reliability was high for complexity (mean length of clauses = .98; subordination = .98; coordination = .97), accuracy (errors per t-unit = .82; the ratio of error free t-units to the total number of t-unit = .81), content (.93), organisation (.86), and overall analytic scoring (.92) in the pilot study. The raters resolved their remaining disagreements through discussion.

3.8. Participants

One hundred and forty participants were recruited on a voluntary basis from among the pool of 364 upper-intermediate EFL learners attending English classes at four major language schools in central Iran. The data collected through the general background questionnaire, adapted from Ong (2010), revealed the following information about the participants. Fifty six (40%) were male and 84 female (60%)
adult EFL learners, whose age ranged from 19 to 35 years ($M = 25$). They had studied English for a relatively long period of time ($M = 4$ years and 3 months) and only a small number of them ($n = 5$) 3.57% spoke a foreign language other than English; they spoke Armenian, French, and Russian.

The participants have not resided in any English speaking countries and did not report any writing test results. The majority of the participants were Persian and spoke Persian with friends and family members ($n = 135$) 96.5%. The vast majority of participants rated the frequency of their own English use in the classroom Frequent ($n = 133$) 95% and out of classroom Not Very Frequent ($n = 136$) 97.14%. The participants ($n = 140$) reported receiving moderate hours of oral instruction per week ($M = 3$ hours) and negligible hours for writing instruction per week ($M = 30$ minutes). They reported that they are most familiar with letter-writing, personal reflection/diaries/journals, descriptive, narrative, argumentative, and exposition types of writing, respectively.

Upper-intermediate learners who shared a similar educational background were recruited for this project, as when investigating the effects of task complexity on CALF, Kuiken and Vedder (2008) proposed that a threshold level should be reached before learners’ can judge the complexity of the tasks. In fact, the Threshold Hypothesis (Cummins, 1979) postulates that to complete a task successfully in a language, a threshold level, a certain level of proficiency, must be reached (see, Schoonen, Hulstijn, & Bossers, 1998; Schoonen et al., 2003). For example, it is suggested that beginners might find both the simple and complex tasks similarly difficult, and thus their writing production might not be influenced in terms of CALF.
Although Kuiken and Vedder’s (2008) findings showed that the effect of task complexity was not related to the level of proficiency, upper-intermediate learners were recruited for this study in line with the predictions of the Threshold Hypothesis (Cummins, 1979) and with prior research that supported this hypothesis (Schoonen, et al., 1998, in reading; Schoonen et al., 2003, in writing).

### 3.9. Context of the Study

In Iran, parallel to the formal English instruction provided by the public schools, private language schools offer English courses for all levels of proficiency and for a wide range of learners from three-year olds to senior citizens. The private schools typically offer intensive and extensive English language courses. The courses are taught based on the principles of the communicative approach and comprise four skills: Listening, speaking, reading, and writing, with a focus on the grammar component (Rahimi & Zhang, 2013; Zhang & Rahimi, 2014). Typically, English courses are taught by near-native English-speaking teachers who hold university degrees (e.g., MA) in teaching English as a foreign language (TEFL). The aim of the programmes is typically to enable students to communicate with others in both oral and written English. Preparation courses for international tests such as the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) are also offered for those who plan to further their education in English medium universities or to immigrate to English speaking countries.

The participants were recruited from private language schools mainly for two reasons. First, public school students typically do not achieve upper-intermediate
level of proficiency because of Iran’s dominantly grammar translation mainstream English language teaching, which focuses on structural properties of English language and reading comprehension (Jahangard, 2007; Hayati & Mashhadi, 2010; Riazi, 2005). Second, wide-ranging proficiency levels of students majoring in English at undergraduate level make establishing homogeneity of groups difficult.

The language schools, where this study was conducted, offer English courses at six proficiency levels (two semesters each): Starter, Levels 1 (Elementary), 2 (low intermediate), 3 (intermediate), 4 (upper-intermediate), and 5 (advanced) with a typical class size of 10 to 15 students. After each proficiency level, students enrol in a Free Conversation course. In the Free Conversation classes, students are provided with the opportunities to discuss the topics covered in previous levels and receive peer and teacher corrective feedback in order to further develop their communicative proficiency. The majority of students at the upper-intermediate level are high-school and university students taking the classes in the evenings as an extracurricular activity (Rahimi & Zhang, 2015).

3.10. Procedure

The upper-intermediate level learners of English as a foreign language (EFL) were recruited to write two tasks, one cognitively simple and one complex, based on Robinson’s (2001a) Cognition Hypothesis. Before the commencement of the study, the participants were informed about the study and told that participation is voluntary, and that they can withdraw from the study at any point without any consequences. Then, they signed the informed consent form. Additionally, prior to the main experiment, Oxford Quick Placement test version 2 (2001) and a pre-writing
test (Appendix C), an argumentative writing test similar to the IELTS (the International English Language Testing System) Academic Writing Task 2, were used to assess learners’ proficiency level and writing ability, respectively. Learners whose placement scores fell into upper-intermediate proficiency range (40-47 out of 60) and achieved good to average writing ability scores (68-85 out of 100) were recruited.

The recruited participants were randomly assigned to three groups: One group for Study 1 (60 students) and two groups for Study 2 (40 participants for each group). For Study 1, five participants did not complete at least one of the writing tasks as instructed, and their writing production did not meet the communicative demands of the task; it was not communicatively adequate (Pallotti, 2009). Therefore, their data were excluded from the main data analysis. To ensure the homogeneity of the groups, after assuring that the assumptions of normality and homogeneity of variance were met, two one-way ANOVAs were run on the three groups’ placement and pre-writing scores. There were no statistically significant differences between the groups regarding their pre-writing, $F(2, 137) = .383, P = .683$, and placement tests, $F(2, 137) = .371, P = .691$.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Pre-writing test</th>
<th>Placement test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Complexity Group</td>
<td>60</td>
<td>73.48</td>
<td>3.35</td>
</tr>
<tr>
<td>Complexity+Planning</td>
<td>40</td>
<td>73.97</td>
<td>3.91</td>
</tr>
<tr>
<td>Complexity+No-Planning</td>
<td>40</td>
<td>74.05</td>
<td>3.48</td>
</tr>
</tbody>
</table>

The main study was conducted in two sessions held on two separate days. The participants for Study 1 completed the individual learner differences (ID) and general
background (GB) questionnaires and the participants of Study 2 GB questionnaire in the first session. The main experiments were administered in the second session as displayed in Figures 3.1 and 3.2. The ID questionnaires should be administered prior to and in as close temporal proximity as possible to the task performance, the outcomes of which are compared with the ID questionnaire findings (Pajares, 2003). However, task-specific questionnaires, which aim at eliciting data on the effect of tasks on ID or on the writing processes employed by the writers should be administered after the task completion and in as close temporal proximity as possible to the task performance. Therefore, the ID questionnaires and post-writing questionnaires were administered prior to and immediately following the task completion, respectively.

In the second session, the learners completed the writing tasks, followed by the task specific questionnaires. Specifically, the participants completed the first task, task difficulty questionnaire, the second writing task, and task difficulty questionnaire again (see Figures 3.1 & 3.2). The tasks were counterbalanced in order to avoid potential practice and fatigue effects (as in Robinson, 2001a & Révész, 2011). That is, half of the learners took the simple task first and the other half of the learners the complex task.
In Study 2, to investigate the synergistic effects of task complexity (+/- reasoning and few/more elements along resource-directing and +/- pre-task planning time along resource-dispersing dimension) on L2 writing production, participants performed the tasks under the pre-task-planning versus no-pre-task-planning conditions. The pre-task planning group received an extra 10 minutes for planning and were instructed to plan their essay before writing. They were provided with a planning sheet to take notes; their planning sheet were removed prior to starting writing to prevent participants from drafting their writing during planning time (as in Adams, Newton, Amani, & Alwi, 2014; Yuan & Ellis, 2003). The no-pre-task planning group were instructed to start writing immediately and to complete their writing tasks in 35 minutes. The 35-minute time frame was set based on the results of the pilot study. To account for the time spent for the completion of each writing task, time-on-task, by each participant, the time spent by each individual learner on each task was recorded.
The independent variable of this study is task complexity in terms of the resource-directing and resource-dispersing dimensions based on Robinson’s (2001a) Cognition Hypothesis, while the writing production measures (CALF, organisation, content, and text quality) compose the dependent variables. The ID variables, measured via self-reported questionnaires, are treated as modulating factors. Level of proficiency was treated as a controlled variable, as only upper-intermediate level learners were recruited for this study.
3.11. Materials

In this section, the research instruments are described. These are: Writing Tasks, Writing Motivational Beliefs Questionnaire, Writing Task Difficulty Questionnaire, Writing Anxiety Questionnaire, and Jacob et al.’s (1981) Analytical Rating Scheme.

3.11.1. Tasks

Similar to Révész’s (2011) oral production study, two experimental tasks, adapted from Watson, DeSanctis, and Poole (1988), were used in this study. The rationales for choosing argumentative decision-making tasks for this study are as follows. First, as Révész has pointed out, prior research studies (Foster & Skehan, 1996; Skehan & Foster, 1997) showed that argumentative tasks, to borrow their terms “decision-making” tasks, generated consistent patterns in measures of accuracy and linguistic complexity. Second, argumentative writing is a typical discourse mode in English courses. Third, prior research has not explored the role of increasing task complexity along the resource-directing and resource-dispersing dimensions in argumentative L2 writing task production.

In line with Révész’s (2011) oral production study, participants were instructed to imagine themselves as a government official in charge of allocating funds for public projects. They were instructed to evaluate the merits of each project and allocate available funds only on merit bases. Each participant wrote about two different sources of funds: Task 1 and Task 2. Task 1, the cognitively simple task, involved allocating $5,000,000 among three competing projects, whereas in Task 2, a cognitively complex task, they were instructed to allocate the second fund, which amounted to $10,000,000, across six competing projects.
The competing projects were all developed for public causes such as improving public transportation, providing free health services for the needy, building cheap housing for the poor, and creating job opportunities (see Appendix A). All the projects were hot topics among the Iranian students and expected to be engaging to the participants. Again, they were instructed to allocate money based on merit only and to provide strong arguments for their prioritising of the projects and for the amount allocated for each project, sufficient to convince the general public. In line with the Robinson’s Triadic Framework along the +/- reasoning and the +/- few elements dimensions, Task 2, which entails decision making for a larger amount of fund and a greater number of projects, is considered more complex, as it requires participants to provide arguments for allocating greater resources across a greater number of competing projects than does the simple task. As Révész (2011) has convincingly observed “in the fields of SLA (Ellis, 2003; Prabhu, 1987; Robinson, 2001a, 2005) and cognitive psychology (e.g., Halford, Cowan, & Andrews, 2007)” completion of tasks that entail more reasoning and more elements are accepted to be more cognitively complex than ones with fewer demands on reasoning and fewer elements.

To establish the complexity of the six-project task compared to that of the three-project task, 10 teachers and 140 participating students judged the complexity of the two tasks and the students rated the difficulty level of two tasks on a 100-point Likert-type scale. Iterative readings (Patton, 2002; Dörnyei, 2007) of teachers and participating students’ reasons for their ratings of the complexity of the two tasks confirmed the validity of the researcher’s operationalisation of the complexity of the (six-project) task compared to that of the simple (three-project) task. Ninety two
percent of the participants in the main study \((n = 129)\), all the participants in the pilot study \((n = 10)\), and all the teachers considered the six-project task to be more difficult than the three-project task.

This finding lends support to prior studies that showed increasing the number of elements and degree of reasoning results in increasing the complexity of the tasks (Prabhu, 1987; Révész, 2011; Robinson, 2001a, 2005). Additionally, the findings of this study extend the previous findings by revealing that when the elements are closely related and deciding about one element affects the other elements (termed elements trade-off in this study), the task becomes more cognitively complex than when there is no such an effect. For instance, the three project task in this study is more complex than giving directions for moving from point A to B and then to C, also a three-element task, but one where there is no elements trade-off effect between A, B, and C.

In summary, the following reasons emerged from the iterative readings of the participants’ reasons for the complexity of the complex versus the simple task: the larger amount of total fund to allocate, the greater number of elements (projects), the requirement for more reasoning to prioritise this number of projects, the competition (trade-off) among the larger number of elements (projects), and the need for more concentration and time and a wider range of vocabulary items. Further, participating students’ ratings of the difficulty of the two tasks on a 100-point Likert-type scale confirmed that the construct of task complexity was operationalised correctly in this study. The students rated the complex task \((M = 350.17 \text{ out of } 400)\) as more difficult than the simple task \((M = 176.02 \text{ out of } 400)\). As the data violated the assumptions of
Paired Samples $t$ test, a Wilcoxon Signed-rank test was run to check the statistical significance of the differences of the participants’ ratings of the two tasks. The result shows that the participants’ ratings of the tasks were statistically significant $z = -10.26, p < .001, r = .88$, and the effect size was large.

### 3.11.2. Motivational Beliefs Questionnaire

The Motivational Beliefs questionnaire, adapted from Troia et al., (2012), is designed to elicit data on self-efficacy, goal orientation, task value, and success attribution in writing. The questionnaire consists of 30 items each scored on an 11-point response scale ranging from 0 (totally disagree) to 100 (totally agree). An 11-point response scale ranging from 0 (totally disagree) to 100 (totally agree) rather than the traditional 5-point response ranging from 1 (strongly disagree) to 5 (strongly agree) was used for each item as prior studies have shown that the former is psychometrically stronger than the latter (Pajares, Hartley, & Valiante, 2001), which corroborates Bandura’s (1997) advice that “including too few steps loses differentiating information because people who use the same response category would differ if intermediate steps were included” (p. 44).

The original questionnaire (Troia et al., 2012) also includes items for collecting data on students’ writing activities; for this study, only items addressing motivational beliefs were used. Table 5 displays the components of the Motivational Beliefs questionnaire. Many of the items on Troia et al.’s (2012) scale were adapted from the scale developed by Pajares, Hartley, and Valiante (2001), Eccles et al. (1989), and Shell et al. (1995). The negatively stated items were reverse scored. The motivational beliefs questionnaire was administered to the participants in the first study in order to
evaluate the effect of interaction between task complexity and motivational beliefs on foreign language writers’ writing production and text quality. The overall Cronbach’s Alpha reliability coefficient for the questionnaire was acceptable ($a = .81$).

### Table 3.2. Questions on Motivational Beliefs Questionnaire

<table>
<thead>
<tr>
<th>Questions on</th>
<th>Sub-category</th>
<th>No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>Efficacy expectations for writing skills and strategies</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Perceived competence for writing tasks</td>
<td>2</td>
</tr>
<tr>
<td>Attribution</td>
<td>Internal attributions (ability and effort)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>External attributions (luck and task ease)</td>
<td>2</td>
</tr>
<tr>
<td>Goal orientation</td>
<td>Mastery goal</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Performance goal</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Avoidance goal</td>
<td>6</td>
</tr>
<tr>
<td>Task value</td>
<td>Task interest/value</td>
<td>5</td>
</tr>
</tbody>
</table>

#### 3.11.3. Writing Task Difficulty Questionnaire

The Task Difficulty questionnaire, adapted from Robinson (2001a), is a task specific questionnaire that was administered immediately after task completion. It consists of questions developed to assess overall task difficulty, level of stress, degree of motivation to complete the writing task, ability to complete the writing tasks, and interest in the writing task. The items of the questionnaire are scored on the 100-point response Likert type scale and are worded as follows. The negatively worded items were reverse coded. For this study, only items eliciting data on task difficulty were used (items: 1, 2, 5, 6), as two other comprehensive questionnaires were employed to elicit data on participants’ motivational beliefs and anxiety. The Cronbach’s Alpha reliability coefficient for this questionnaire was above the acceptable level ($a = .77$).

1. I thought this writing task was easy;
2. I thought this writing task was difficult;
3. I was relaxed writing this essay;
4. I was annoyed writing this essay;
5. I didn’t do well on this writing task;
6. I did well in this writing task;
7. This writing task was not interesting;
8. This writing task was interesting;
9. I don’t want to do more writing tasks like this;
10. I want to do more writing tasks like this.

3.11.4. Writing Anxiety Questionnaire

Three subscales, Somatic Anxiety (items: 2, 7, 9, 13, 15, 18, 23), Avoidance Behaviour (items: 4, 6, 12, 14, 19, 22, 27), and Cognitive Anxiety (items: 1, 3, 8, 10, 17, 21, 24, 26), consisting a total of 22 items compose the Writing Anxiety Questionnaire, which is adapted from Cheng (2004). Seven negatively worded items were reverse coded. A higher score obtained on the subscales and the total scale of the writing anxiety questionnaire show a higher level of L2 writing anxiety.

An 11-point response scale ranging from 0 (totally disagree) to 100 (totally agree) were used for each item instead of the original 5-point response scale ranging from 1 (strongly disagree) to 5 (strongly agree), as previous research has demonstrated that a scale with a 0-100 response format is psychometrically stronger than one with a traditional Likert format. For instance, in Pajares, Hartley, and Valiante’s (2001) study, the 0-100 scale writing skills self-efficacy predicted both middle school students’ grade point average in language arts and teachers’ ratings of their students’ writing competence, whereas the traditional Likert scale rating did not. To illustrate,
the variance linked exclusively with the Likert scale was insignificant for both grade point average and each teacher’s ratings, whereas 37% of the variance in grade point average and 28% of the variance in teachers’ ratings was linked exclusively with the 0-100 response format scale. This finding clearly substantiates Bandura’s (1997) advice that “including too few steps loses differentiating information because people who use the same response category would differ if intermediate steps were included” (p. 44). Since a 0-100 Likert type response scale is neither more difficult nor longer than the traditional Likert format response scale, the former was administered to the participants of Study 1, as it enhances the predictive utility of the questionnaire (Troia et al., 2012). The Cronbach’s Alpha reliability coefficients for the three subscales, Somatic Anxiety (.78), Avoidance Behaviour (.76), and Cognitive Anxiety (.82) were acceptable.

3.11.5. Jacob et al.’s (1981) Analytical Rating Scheme

Jacob et al.’s (1981) Analytical Rating Scheme, a well-established and widely-used composition rating scheme (e.g., Cumming, 1989; Ong & Zhang, 2013; Rao, 2007, Sasaki, 2002), was used to quantify components of the learners’ written text quality. Jacob et al.’s (1981) Analytical Rating Scheme is a multiple trait rating scheme (Hamp-Lyons, 1991) which allows the assessment of written text quality in terms of five components: Content (30 marks), organisation (20 marks), language (25 marks), vocabulary (20 marks), and mechanics (5 marks). Jacob et al.’s (1981) Analytical Rating Scheme, rather than a holistic measure, was used for rating the components of the learners’ argumentative writing for two main reasons. First, the assessment of various components of the writing typically results in more reliable scores than the holistic ones (Hughes, 1989, 2003), although holistic schemes are less labour-
intensive and less time-consuming (Ong & Zhang, 2013). Second, to establish whether there is a trade-off between higher-order writing skills, separate scores for the content and organisation components of the learners’ writing production are needed.

3.12. Independent Variables

Based on Robinson’s Triadic Componential Framework (Robinson, 2001a, 2001b, 2003, 2005, 2007, 2011a, 2011b), the independent variable of this study is task complexity along the resources-directing and resource-dispersing dimensions. The writing tasks are manipulated along the resource-directing dimension by increasing the degree of reasoning and the number of elements. That is, the cognitively complex argumentative writing task requires higher level of reasoning and has more elements relative to the cognitively simple task.

Also, the complexity of the two writing tasks is further manipulated along the resource-dispersing dimension by providing a 10 minute pre-task-planning time to one group to plan the content, organisation, and language, whereas the other group had no-pre-task-planning time and started writing immediately. This condition can be defined as no-pre-task-planning condition or pressured online planning condition (Ellis, 2005), as the participants write with no-pre-task-planning time, but they might plan and write as they proceed in the recursive process of writing.

3.13. Dependent Variables

The dependent variables under investigation in this study are task difficulty and writing production. L2 writing production measures comprise syntactic complexity,

In this study, in addition to general task production measures, measures of organisation (cohesion and coherence), content, and text quality were also used by employing Jacob et al.’s (1981) analytical rating scheme. The rationale for adding these three measures to the general measures of complexity, accuracy, and fluency is the conviction that there might be a trade-off between these higher-order measures rather than just between the general measures of writing production as stated in prior research:

Additionally, there are other aspects of performance worth considering. In our experiment, no attention was paid to the actual content or the argumentative force of the text. No assessment was made of the effect of task complexity on the use of higher-order writing skills such as the cohesion and coherence of the text. It may be the case that these aspects of performance are also affected by an increase in task complexity. Moreover, it could be that the extra attention paid to accuracy in the more complex task condition is taken away from these higher-order processes. As a consequence, these aspects certainly deserve to be studied further (Kuiken & Vedder, 2008, p. 28).
3.14. Writing Production Measures

The measures used to detect the impacts of task complexity on the dimensions of L2 writing production are described in this section. Each measure of L2 writing production, namely, syntactic complexity, accuracy, lexical complexity, and fluency are defined, and the rationales for adopting each measure are discussed.

3.14.1. Syntactic Complexity

Syntactic complexity refers to the extent to which learners can use the language “closer to the cutting edge of inter-language development” (Foster & Skehan, 1996, p. 304). In this study, a multiple conceptualisation of complexity (Norris & Ortega, 2009) is adopted to capture the effects of task complexity on the dimensions of L2 syntactic complexity. The reasoning is that task complexity might influence the dimensions of syntactic complexity of learners’ writing production differently, depending on the students’ level of proficiency.

In an attempt to capture task complexity effects on all dimensions of syntactic complexity, which are shown to be the distinctive features of students’ interlanguage development at different proficiency levels, and to avoid using redundant measures, which might lead to multicollinearity effect (Norris & Ortega, 2009), three measures of complexity are employed in this study. These are: Mean Length of Clauses (MLC) as a measure of sub-clausal complexity, an advanced proficiency level complexity measure, subordination (SUB) as a measure of clausal complexity, an upper-intermediate level complexity measure, and phrasal coordination as a beginner proficiency level complexity measure. Hunt’s (1966, p. 735) definition of t-unit,
which refers to t-unit as “one main clause plus whatever subordinate clauses happen to be attached to or embedded within it”, is adopted for this study.

T-unit rather than C-unit or AS-Unit, the Analysis of Speech Unit, (Foster, Tonkyn, & Wigglesworth, 2000) was adopted due to monologic nature of argumentative writing production. MLC, subordination, and phrasal coordination were calculated by dividing the total number of words by the total number of clauses, the total number of subordinate clauses by the total number of clauses, and the total number of coordinate phrases by the total number of t-units, respectively.

3.14.2. Accuracy

In this study, Skehan’s (1996) definition of accuracy is adopted. Accuracy refers to the extent the produced language abides by the rules of the given language and is measured by general measures such as error-free clauses (as in Crookes, 1989) and specific measures such as correctness of regular simple past tense (e.g., Ellis, 1987). The accuracy of the learners’ argumentative writing was measured by calculating the ratio of error-free t-units to the total number of t-units (as in Arent, 2003; Storch, 2009). An error-free t-unit is defined, for the purpose of this study, as a t-unit that is free of any lexical, grammatical, or morphological errors. All errors in spelling and punctuation were overlooked to reduce the possible over-estimation of errors owing to participants’ indistinct handwriting (as in Storch, 2009). Additionally, as the ratio of error-free t-units to the total number of t-units does not account for the number of errors per t-unit, which might be affected by task complexity, the number of errors per t-unit was also measured in this study. Polio and Shea’s (2014) guidelines was used in coding the errors.
3.14.3. Lexical Diversity and Academic Vocabulary Use

For lexical diversity, values of D (Malvern & Richards, 2002) were calculated by using the computer programme vocd in CLAN (MacWhinney, 2000). Other measures of lexical diversity that are calculated by the type–token ratio (TTR), the ratio of different words (Types) to the total number of words (Tokens), are sensitive to produced sample (written text) length, as MacWhinney (2002) has stated:

Unfortunately, such measures, including mathematical transformations of the TTR such as Root TTR, are functions of the number of tokens in the transcript or language sample — samples containing larger numbers of tokens give lower values for TTR and vice versa (Richards & Malvern, 1997a). This problem has distorted research findings (Richards & Malvern, 1997b). Previous attempts to overcome the problem, for example by standardizing the number of tokens to be analyzed from each child, have failed to ensure that measures are comparable across researchers who use different baselines of tokens, and inevitably waste data in reducing analyses to the size of the smallest sample (MacWhinney, 2000, p. 139).

To measure advanced academic vocabulary use, Range 32 software (Nation, 2007) was used, and the learners’ use of words on the third word list, which are believed to appear in university texts, was calculated as a measure of lexical sophistication.

3.14.4. Fluency

In line with Skehan’s (1996) definition of fluency, which equates fluency with the extent to which language learners can use their interlanguage to communicate their intended meaning in real time, fluency was measured by calculating the total number of words per minute (as in Kellogg, 1990; Ong & Zhang, 2010; Storch, 2009). The total number of words produced by each participant in the argumentative writing tasks was divided by the total number of minutes spent for writing by each participant on the same writing task. To record the time each participant spend on
each writing task, participants started writing at the same time, and their finishing time was recorded by the research assistant.

3.15. Reliability of Data Coding and Rating

All data except lexical diversity and academic vocabulary use were coded and rated by the research assistant, and a randomly selected 40 percent of the collected data was re-coded and re-rated by the researcher. The Spearman’s rho correlation for inter-coder and inter-rater reliability was high for complexity (mean length of clauses = .96; subordination = .98; coordination = .98), accuracy (the number of errors per t-unit = .84; the ratio of error-free-t-units to the total number of t-units = .86), fluency (.99), content (.91), organisation (.93), overall analytic scoring (.95) for the main study, and for overall analytic scoring for the pre-writing test (.89). Additionally, to investigate intra-coder and intra-rater reliability, the research assistant re-coded and re-rated 40 percent of the data a month apart. The Cronbach’s Alpha for intra-coder reliability for complexity (mean length of clauses = .99; subordination = .97; coordination = .98), accuracy (the number of errors per t-unit = .92; the ratio of error-free-t-units to the total number of t-units = .93), fluency (.99), content (.98), organisation (.97), and overall analytic scoring (.97) was well above the acceptable level.

3.16. Statistical Analyses

Statistical analyses were chosen according to the respective research questions and the characteristics of the obtained scores. First, the assumptions of the statistical tests were checked via standard diagnostic tests and procedures. For both studies, repeated measures MANOVAs were not conducted, as the assumptions of MANOVA were
not met. For Study 1, to address the impact of task complexity on L2 writing production measures, paired samples $t$ tests and Wilcoxon Signed Ranks tests were conducted depending on the normality of the distribution of the obtained scores (as in Révész, 2011). To address Skehan’s (2009) concern that the positive experimental effects do not ascertain simultaneous improvements in the dimensions of L2 production, in cases where significant positive experimental effects were found for more than one dimension of CALF, follow-up correlation and frequency analyses were conducted to identify the number of individuals who experienced simultaneous improvements and the trade-off effect. To explore the modulating role of ID variables, correlations were computed between the ID variables and writing production measures (as in Révész, 2011).

For the second study, mixed within and between groups repeated measures ANOVAs were conducted to examine the main effects for planning conditions (present vs. absent) and task complexity, and the interaction effect between planning and task complexity on the measures of writing production (as in Révész, Michel, & Gilabert, 2015). As the main objective of Study 2 was comparing the learners’ production across four conditions, further follow-up paired samples $t$ tests or Wilcoxon Signed Ranks tests for within groups complexity effects across the simple and complex tasks and independent samples $t$ test or Mann-Whitney $U$ test for between groups across planning conditions were run to further explore the effects of task complexity and planning conditions on writing production measures. For the scores that did not meet the assumptions of ANOVA, only within and between groups analyses were conducted. Follow-up procedure was applied even for the measures that ANOVA did not yield significant effect (as in Yuan, 2001). However, to account for the possible
inflated Type I errors, the alpha level for each significant effect was checked. Fortunately, the alpha level for almost all significant results was very small, around $p < .001$, although initially an alpha level of $p < .05$ was set for all tests. To measure effect sizes, Cohen’s $d$ was calculated for the $t$ tests and $r$ was calculated for the Wilcoxon and Mann-Whitney tests; $d$ values of .20, .50, and .80 and $r$ values of .10, .30, and .50 were considered small, medium, and large, respectively, according to Cohen’s (1992) guidelines. For statistical power analysis G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) was used. The sample sizes were adequate to find medium effect sizes for all variables with an $\alpha = .05$ and power = .80.
Chapter 4.

Results for Study 1

4.1. Chapter Overview

The results of Study 1, the effect of cognitive task complexity on, and the modulating role of motivation and anxiety in, L2 writing syntactic and lexical complexity, accuracy, fluency, content, organisation, and overall quality, are presented in this chapter. A summary of the findings closes the chapter.

4.2. Exploring Assumptions for Statistical Tests

To examine the effect of task complexity on English-as-a-foreign-language (EFL) learners’ writing production, first, the assumption of normality was checked. The assumptions of repeated measures MANOVA were not met. As a repeated measure design was adopted to investigate the effect of task complexity on the dimensions of EFL learners’ writing production, the difference between each pair of score was computed, and the assumption of normally distributed difference scores was examined via Shapiro-Wilk’s test, the visual inspection of histograms, normal Q-Q plots and box plots, and the Skewness and Kurtosis of each computed score (see Table 4.1).
Table 4.1. Normality Test Results

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>N</th>
<th>Shapiro-Wilk Statistic</th>
<th>Sig.</th>
<th>Kurtosis Statistic</th>
<th>Std. Error</th>
<th>Skewness Statistic</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>F(S-C)</td>
<td>55</td>
<td>.95 (.02*)</td>
<td></td>
<td>2.8</td>
<td>.63</td>
<td>.40</td>
<td>.32</td>
</tr>
<tr>
<td>MLC(S-C)</td>
<td>55</td>
<td>.98 (.67)</td>
<td></td>
<td>-.16</td>
<td>.63</td>
<td>.07</td>
<td>.32</td>
</tr>
<tr>
<td>SUB(S-C)</td>
<td>55</td>
<td>.94 (.02*)</td>
<td></td>
<td>1.05</td>
<td>.63</td>
<td>-.10</td>
<td>.32</td>
</tr>
<tr>
<td>COR(S-C)</td>
<td>55</td>
<td>.42 (.000*)</td>
<td></td>
<td>8.52</td>
<td>.63</td>
<td>-.296</td>
<td>.32</td>
</tr>
<tr>
<td>ERPT(S-C)</td>
<td>55</td>
<td>.93 (.008)</td>
<td></td>
<td>-.04</td>
<td>.63</td>
<td>-.71</td>
<td>.32</td>
</tr>
<tr>
<td>ERFRT(S-C)</td>
<td>55</td>
<td>.97 (.03*)</td>
<td></td>
<td>.90</td>
<td>.63</td>
<td>-.83</td>
<td>.32</td>
</tr>
<tr>
<td>LD(S-C)</td>
<td>55</td>
<td>.97 (.19)</td>
<td></td>
<td>.58</td>
<td>.63</td>
<td>-.56</td>
<td>.32</td>
</tr>
<tr>
<td>AWU(S-C)</td>
<td>55</td>
<td>.93 (.003)</td>
<td></td>
<td>2.62</td>
<td>.63</td>
<td>1.08</td>
<td>.32</td>
</tr>
<tr>
<td>CON(S-C)</td>
<td>55</td>
<td>.92 (.001*)</td>
<td></td>
<td>.23</td>
<td>.63</td>
<td>-.63</td>
<td>.32</td>
</tr>
<tr>
<td>ORQ(S-C)</td>
<td>55</td>
<td>.90 (.000*)</td>
<td></td>
<td>.49</td>
<td>.63</td>
<td>-.08</td>
<td>.32</td>
</tr>
<tr>
<td>WQ(S-C)</td>
<td>55</td>
<td>.97 (.35)</td>
<td></td>
<td>1.08</td>
<td>.63</td>
<td>.08</td>
<td>.32</td>
</tr>
</tbody>
</table>

Notes: (S-C) = (the score on the simple task – the score on the complex task), F = fluency, MLC = mean length of clauses, SUB = subordinate clauses per clause, COR = phrasal coordination per t-unit, ERPT = the number of errors per t-unit, ERFRT = error-free-t-units per t-unit, LD = lexical diversity, AWU = academic vocabulary use, CON = content, ORQ = organisation, WQ = writing quality, N = number of participants, Sig. = significance, Std. = standard.

Paired Samples t tests or Wilcoxon Signed Ranks tests were conducted to test the task complexity effects on writing production measures based on the normality test results and the visual inspection of histograms and distribution plots. For MLC (mean length of clauses), SUB (subordinate clauses per clause), ERPT (the number of errors per t-unit), ERFRT (the ratio of error-free-t-units to the total number of t-unit), LD (lexical diversity), and writing quality, paired samples t tests were run and Wilcoxon Signed Ranks tests were run for the rest of the measures.

4.3. Task Complexity Effects on Dimensions of Syntactic Complexity (RQ1)

As I discussed in the previous chapter, Norris and Ortega (2009) have proposed a multidimensional conceptualisation of L2 syntactic complexity and recommended employing multiple measures to capture the development of complexity in learners’ interlanguage. They convincingly argue that different dimensions of complexity develop at different levels of interlanguage development. They also suggest using one measure for each dimension of the complexity to avoid potential
multicollinearity effect. In this study, Norris and Ortega’s (2009) conceptualisation of complexity is adopted, and three measures of complexity, one for each dimension, are employed: Mean Length of Clauses (MLC) as a measure of sub-clausal complexity, subordination (subordinate clauses per clause) as a measure of clausal complexity, and phrasal coordination as a beginner level complexity measure.

As regards the effects of task complexity on the dimensions of complexity, increasing complexity along the resource-directing dimension did not have a significant impact on the learners’ performance in terms of mean length of clauses: Simple task ($M = 9.90, SD = 1.88$) versus complex task ($M = 10.05, SD = 1.74$), $t(54) = .79, p = .43, d = .10$, and phrasal coordination: Simple task ($M = .28, SD = .18$) versus complex task ($M = .27, SD = .20$), $z = -1.11, p = .26, r = .10$. However, there was a significant positive effect on the number of subordinate clauses per clause. The learners wrote significantly more subordinate clauses per clause in the complex task performance ($M = .37, SD = .07$) than in its simple version ($M = .31, SD = .07$), $t = 5.87, p < .001, d = .79$ (see Table 4.2).

<table>
<thead>
<tr>
<th>Tasks</th>
<th>MLC</th>
<th>COR</th>
<th>SUB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Simple</td>
<td>9.90</td>
<td>1.88</td>
<td>.28</td>
</tr>
<tr>
<td>Complex</td>
<td>10.05</td>
<td>1.74</td>
<td>.27</td>
</tr>
<tr>
<td>Sig/Effect</td>
<td>$p = .43, t = .79, d = .10$</td>
<td>$p = .26, r = .10$</td>
<td>$p &lt; .001, t = 5.87, d = .79$</td>
</tr>
</tbody>
</table>

*Note. Sig = significance; Effect = effect size; Simple = simple task; Complex = complex task*

4.4. Task Complexity Effects on Accuracy (RQ1)

Writing accuracy was operationalised as the extent to which the learners’ writing production abides by the rules of English (Skehan, 1996) and was investigated by using two measures that targeted the absence and presence of the errors in learners’
writing production. The absence of errors was measured by calculating the ratio of error-free t-units to the total number of t-units (as in Arent, 2003; Storch, 2009) and the presence of errors by the number of errors per t-unit. In calculating error-free t-units and errors per t-unit all lexical, grammatical, and morphological errors were considered, but all spelling and punctuation errors were overlooked to decrease the possible over-estimation of errors (as in Storch, 2009). For error coding, Polio and Shea’s (2014) guidelines were used. Paired Samples \( t \) tests were run to measure the effects of increasing task complexity on the accuracy of EFL learners’ writing production.

### Table 4.3. Descriptive Statistics for Measures of Writing Accuracy on the Simple versus Complex Tasks \((N = 55)\)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>ERFT/TU</th>
<th>ER/TU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Simple</td>
<td>.33</td>
<td>.18</td>
</tr>
<tr>
<td>Complex</td>
<td>.28</td>
<td>.15</td>
</tr>
<tr>
<td>Sig/Effect</td>
<td>( p = .047 ), ( d = .27 )</td>
<td>( p &lt; .001 ), ( d = .66 )</td>
</tr>
</tbody>
</table>

*Note. Sig = significance; Effect = effect size; Simple = simple task; Complex = complex task*

As shown in Table 4.3, increasing task complexity had significant negative consequences on EFL learners’ writing accuracy as measured by both error-free-t-units per t-unit: Simple task \((M = .33, SD = .18)\) versus complex task \((M = .28, SD = .15)\), \( t(54) = 2.03, p = .047, d = .27 \) and by the number of errors per t-unit: Simple task \((M = 1.25, SD = .53)\) versus complex task \((M = 1.62, SD = .70)\), \( t(54) = -4.98, p < .001, d = .66 \).

### 4.5. Task Complexity Effects on Lexical Diversity (LD) and Academic Vocabulary Use (AWU; RQ1)

Values of D (Malvern & Richards, 2002) were calculated as a measure of lexical diversity by using the computer programme vocd in CLAN (MacWhinney, 2000). This measure was selected because other measures of lexical diversity that are
calculated by the type–token ratio (TTR), the ratio of different words (Types) to the total number of words (Tokens), are sensitive to produced sample (written text) length (MacWhinney, 2002). As the normality of assumption was met for LD, a paired samples t test was run to measure the task complexity effect on LD. As displayed in Table 4.4, the learners’ performance on the complex version of the task ($M = 78.96, SD = 14.05$) did not significantly differ from that on the simple version ($M = 77.04, SD = 13.18$), $t(54) = -1.25, p = .21, d = .17$.

Range 32 software (Nation, 2007) was used to measure learners’ AWU. As the assumption of normally distributed difference score was found to be violated, Wilcoxon Signed Rank test was performed to test the effect of task complexity on AWU. The learners’ AWU improved significantly on the complex ($M = 1.70, SD = .82$) versus the simple version of the task ($M = .94, SD = .87$) as a function of increasing task complexity, $z = -5.356, p < .001, r = .72$, with a large effect size.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple task</th>
<th>Complex task</th>
<th>Sig/Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>D value</td>
<td>77.04 13.18</td>
<td>78.96 14.05</td>
<td>$p = .21, d = .17$</td>
</tr>
<tr>
<td>AWU</td>
<td>.94  .87</td>
<td>1.70  .82</td>
<td>$p &lt; .001, r = .72$</td>
</tr>
</tbody>
</table>

Note. Sig = significance; Effect = effect size

As the experimental effect on both SUB and AWU was significant, the correlation and frequency analyses were run to identify the individual level differences. The results for the correlation between SUB and AWU in complex task performance was not significant, $r = .15, p = .27$. However, the frequency analysis showed simultaneous improvements, the trade-off effect, improvement in one dimension and no change in the other, and simultaneous decreases in 67.3, 10.9, 18.2, and 3.6.
percent of the cases, respectively. Therefore, the simultaneous improvement, not the trade-off effect, was the norm at the individual level.

4.6. Effect of Task Complexity on Fluency (RQ1)

In line with Skehan’s (1996) definition, fluency was operationalised as the extent to which language learners use their interlanguage to communicate their intended meaning in real time. Fluency was measured only by calculating the total number of words per minute (as in Kellogg, 1990; Ong & Zhang, 2010; Storch, 2009), as using other types of fluency measures, which are typically used in oral production, was not suitable for measuring writing fluency. To record ‘time-on-task’ for each writing task, participants started writing at the same time, and their finishing time was recorded by the research assistant. As the assumption of normality was violated for fluency, Wilcoxon Signed Ranks test was performed to examine the task complexity effect on fluency.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple task</th>
<th>Complex task</th>
<th>Sig/Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words per minute</td>
<td>9.30/2.77</td>
<td>7.99/2.29</td>
<td>$p &lt; .001, r = .51$</td>
</tr>
</tbody>
</table>

As displayed in Table 4.5, the students performed significantly more fluently on the simple ($M = 9.30, SD = 2.77$) than on the complex version of the task ($M = 7.99, SD = 2.29$), $z = -3.80, p < .001, r = .51$. Increasing task complexity had a deteriorating effect on fluency, with a large effect size ($r = .51$).
4.7. Task complexity Effects on Content, Organisation, and Overall Writing Quality (RQ1)

To investigate the effects of task complexity on task content, organisation, and overall writing quality in addition to CALF, separate scores were needed for these components of the learners’ writing production. Jacob et al.’s (1981) Analytical Rating Scheme was employed to quantify these components. As displayed in Table 4.6, increasing task complexity prompted the learners to perform significantly better on the complex version of the task than on its simple version in terms of content: Simple ($M = 22.69, SD = 1.66$) versus the complex task ($M = 23.09, SD = 1.87$), $z = -2.29, p = .022, r = .31$; organisation: Simple ($M = 15.19, SD = 1.38$) versus the complex task ($M = 15.62, SD = 1.49$), $z = -2.71, p = .007, r = .36$; and overall writing quality: Simple ($M = 74.71, SD = 5.29$) versus the complex task ($M = 76.14, SD = 6.66$), $t(54) = -2.66, p = .010, d = .36$.

Table 4.6. Measures of Content, Organisation, and Overall Quality on the Simple versus Complex Tasks ($N = 55$)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Content</th>
<th>Organisation</th>
<th>Overall Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Simple</td>
<td>22.69</td>
<td>1.66</td>
<td>15.19</td>
</tr>
<tr>
<td>Complex</td>
<td>23.09</td>
<td>1.87</td>
<td>15.62</td>
</tr>
<tr>
<td>Sig/Effect</td>
<td>$p &lt; .05$</td>
<td>$r = .31$</td>
<td>$p &lt; .05$</td>
</tr>
</tbody>
</table>

Note. Sig = significance; Effect = effect size; Simple = simple task; Complex = complex task

4.8. Effects of Task Complexity on Task Difficulty (RQ2)

In this study, to establish the complexity of the complex (six-project) task compared to that of the simple (three-project) task, 10 teachers and participating students judged the complexity of the two tasks, and the students rated the difficulty level of the two tasks on a 100-point Likert-type scale. The result of students’ rating of the two tasks and iterative readings (Patton, 2002) of the teachers’ and participating students’ reasons for the complexity level of the two tasks confirmed the validity of
the researcher’s operationalisation of the higher complexity of the complex (six-project) task when compared to that of the simple (three-project) task. Ninety two percent of the participants in the main study \((n = 129)\), all the participants in the pilot study \((n = 10)\), and all the teachers found the six-project task more difficult than the three-project task.

This finding lends support to prior studies that showed that increasing the number of elements and the degree of reasoning results in increasing the complexity of tasks (Prabhu, 1987; Révész, 2011; Robinson, 2001a, 2005). Additionally, the findings of this study extend the previous findings by revealing that when the elements (projects) are closely related (e.g., are all at a similar level of worthiness for funding) and deciding for one element affects the other elements (termed elements trade-off in this study), the task becomes more cognitively complex than when there is no such an effect. For instance, the three project task in this study is more complex than a task that requires giving directions for moving from point A to B and then to C, also a three-element task, as there is no elements trade-off effect between the elements, A, B, and C. The following emerged from the iterative readings of the participants’ reasons for the complexity of the six-project versus the three-project task: the larger amount of total fund to allocate, the greater number of elements (projects), the requirement for more reasoning to prioritise this number of projects, the competition (trade-off) among the larger number of elements (projects), and the need for more concentration and time and a wider range of vocabulary items.

Furthermore, participating students’ ratings of the difficulty of the two tasks on a 100-point Likert-type scale confirmed that the construct of task complexity was
operationalised appropriately in this study. The students’ rated the complex task \((M = 350.17, SD = 33.93)\) significantly more difficult than the simple task \((M = 176.02, SD = 40.11)\), \(z = -10.26, p < 0.001\). The results were associated with a large effect size, \(r = .88\).

**Table 4.7. Students’ Ratings of the Difficulty of the Simple versus Complex Tasks \((N = 135)\)**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple task</th>
<th>Complex task</th>
<th>Sig/Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M)</td>
<td>(SD)</td>
<td>(M)</td>
</tr>
<tr>
<td>Rating Scale</td>
<td>176.02</td>
<td>40.11</td>
<td>350.17</td>
</tr>
</tbody>
</table>

### 4.9. Modulating Role of Affective Variables in L2 Writing

Before computing correlations between affective variables and writing production measures, the assumption of normality for each variable and measure was checked and appropriate statistical analysis was chosen accordingly.

#### 4.9.1. Modulating Role of Writing Anxiety (RQ3)

To address the third research question, which seeks to answer the extent to which multiple dimensions of anxiety correlate with the dimensions of L2 learners’ writing production, multiple correlations were performed. The only significant correlations found between anxiety and the dimensions of L2 learners’ writing production were a moderate negative relationship between avoidance behaviour and mean length of t-units (MLT), \(r = -.28, p < .05\), as a general measure of writing syntactic complexity, and between avoidance behaviour and AWU, \(r = -.27, p < .05\), as a measure of advanced vocabulary use. Interestingly these significant relationships were both found in the complex task performance (see Tables 4.8, 4.9, 4.10).
4.9.2. Modulating Role of Writing Motivational Beliefs (RQ4)

To answer the fourth research question, multiple correlations were run to examine the extent to which multiple dimensions of motivational belief correlate with dimensions of EFL learners’ writing production. As shown in Tables 4.11, 4.12, and 4.13, there is a significant moderate positive correlation between each of mastery goal, self-efficacy, task value, and attribution internal and mean length of t-units in the complex task, with \( r = .34, .35, .30, \) and \( .28, p < .05 \), respectively (Table 4.11). Maladaptive approaches to learning, namely, avoidance goal \( (r = -.33) \), performance goal \( (r = -.35) \), and external attribution \( (r = -.36) \), each had a significant moderate negative
correlation with MLT. There is also a moderate positive correlation between each of
task value and self-efficacy and the ratio of errors to the total number of t-units in the
complex task, with \( r = .33 \) and \( .35, p < .05 \), respectively (Table 4.12). Further, the
correlation between mastery goal and content reached moderate positive significance
in the complex task, \( r = .31, p < .05 \) (Table 4.13). No significant correlation was
found between motivation and each dimension of L2 writing production in the simple
task. In all cases, positive and significant relationships were found between
dimensions of motivational beliefs and L2 writing production measures in the
complex task performance.

**Table 4.11.** Correlations of Dimensions of Writing Motivation with Measures of
Writing Complexity on the Simple versus Complex Tasks \((N = 55)\)

<table>
<thead>
<tr>
<th>Motivation</th>
<th>MLT</th>
<th>MLC</th>
<th>SUB</th>
<th>COR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sim</td>
<td>Com</td>
<td>Sim</td>
<td>Com</td>
</tr>
<tr>
<td>MG</td>
<td>.16</td>
<td>.34*</td>
<td>.01</td>
<td>.06</td>
</tr>
<tr>
<td>AG</td>
<td>.08</td>
<td>-.33*</td>
<td>.12</td>
<td>-.09</td>
</tr>
<tr>
<td>PG</td>
<td>-.09</td>
<td>-.35*</td>
<td>-.09</td>
<td>.11</td>
</tr>
<tr>
<td>SE</td>
<td>.04</td>
<td>.35*</td>
<td>-.009</td>
<td>.10</td>
</tr>
<tr>
<td>TV</td>
<td>.14</td>
<td>.30*</td>
<td>.12</td>
<td>.21</td>
</tr>
<tr>
<td>ATI</td>
<td>.03</td>
<td>.28*</td>
<td>.22</td>
<td>.21</td>
</tr>
<tr>
<td>ATE</td>
<td>-.008</td>
<td>-.36*</td>
<td>-.03</td>
<td>.09</td>
</tr>
</tbody>
</table>

*Note. MG = Mastery Goal; AG = Avoidance Goal; PG = Performance Goal; SE = Self-efficacy; TV = Task Value; ATI = Attribution Internal; ATE = Attribution External*

**Table 4.12.** Correlations of Dimensions of Writing Motivation with Measures of
Accuracy, Lexical Diversity, Academic Vocabulary Use, and Fluency
on the Simple versus Complex Tasks \((N = 55)\)

<table>
<thead>
<tr>
<th>Motivation</th>
<th>ERTU</th>
<th>ERFTU</th>
<th>D Value</th>
<th>AWU</th>
<th>Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sim</td>
<td>Com</td>
<td>Sim</td>
<td>Com</td>
<td>Sim</td>
</tr>
<tr>
<td>MG</td>
<td>.13</td>
<td>-.10</td>
<td>.01</td>
<td>.08</td>
<td>.13</td>
</tr>
<tr>
<td>AG</td>
<td>-.01</td>
<td>-.12</td>
<td>.05</td>
<td>-.06</td>
<td>.06</td>
</tr>
<tr>
<td>PG</td>
<td>.07</td>
<td>.04</td>
<td>-.05</td>
<td>-.01</td>
<td>.11</td>
</tr>
<tr>
<td>SE</td>
<td>.25</td>
<td>.33*</td>
<td>.21</td>
<td>.23</td>
<td>.01</td>
</tr>
<tr>
<td>TV</td>
<td>.20</td>
<td>.35*</td>
<td>.11</td>
<td>.22</td>
<td>.13</td>
</tr>
<tr>
<td>ATI</td>
<td>.20</td>
<td>.24</td>
<td>.13</td>
<td>.06</td>
<td>.09</td>
</tr>
<tr>
<td>ATE</td>
<td>.17</td>
<td>.25</td>
<td>-.10</td>
<td>-.16</td>
<td>-.14</td>
</tr>
</tbody>
</table>

*Note. Sim = Simple, Com = Complex, MG = Mastery Goal; AG = Avoidance Goal; PG = Performance Goal; SE = Self-efficacy; TV = Task Value; ATI = Attribution Internal; ATE = Attribution External*
Table 4.13. Correlations of Dimensions of Writing Motivation with Content, Organisation, and Overall Quality on the Simple versus Complex Tasks ($N = 55$)

<table>
<thead>
<tr>
<th>Motivation</th>
<th>Content</th>
<th>Organisation</th>
<th>Overall Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Complex</td>
<td>Simple</td>
</tr>
<tr>
<td>MG</td>
<td>.06</td>
<td>.31*</td>
<td>-.08</td>
</tr>
<tr>
<td>AG</td>
<td>.13</td>
<td>.10</td>
<td>.07</td>
</tr>
<tr>
<td>PG</td>
<td>-.01</td>
<td>-.08</td>
<td>-.01</td>
</tr>
<tr>
<td>TV</td>
<td>.21</td>
<td>.23</td>
<td>.14</td>
</tr>
<tr>
<td>ATI</td>
<td>.10</td>
<td>.17</td>
<td>.17</td>
</tr>
<tr>
<td>ATE</td>
<td>-.03</td>
<td>-.18</td>
<td>-.02</td>
</tr>
</tbody>
</table>

Note. MG = Mastery Goal; AG = Avoidance Goal; PG = Performance Goal; SE = Self-efficacy; TV = Task Value; ATI = Attribution Internal; ATE = Attribution External

4.10. Summary of the findings for Study 1

4.10.1. Effects on Dimensions of Syntactic Complexity

Increasing task complexity effected a significant change in only one measure of syntactic complexity, namely, SUB. The effect on other measures (i.e., MLC & PCOR) was not significant. As regards the mediating role of affective variables, adaptive approaches to learning, namely, mastery goal, self-efficacy, task value, and attribution internal each had a moderate positive correlation with MLT in the complex task performance. Maladaptive approaches to learning, namely, avoidance goal, performance goal, and external attribution and one dimension of anxiety, avoidance behaviour, had a significant moderate negative association with MLT in the complex task performance. No significant association was found between either adaptive approaches to learning and other dimensions of syntactic complexity or maladaptive approaches to learning and other dimensions of syntactic complexity in L2 writing production in either the simple or complex task performance.
4.10.2. Effects on Accuracy

Increasing task complexity had a significant adverse impact on accuracy, as measured by both the number of errors per t-unit and the ratio of error-free t-units to the total number of t-units. With regard to the mediating role of anxiety and motivation, there was no significant correlation between the dimensions of writing anxiety and two measures of accuracy in either the simple or complex task performance. Similarly, the correlation between the dimensions of writing motivation and error-free t-units per t-unit was not significant. Conversely, there was a significant moderate positive association between self-efficacy and the number of errors per t-unit, and between task value and the number of errors per t-unit in the complex task performance.

4.10.3. Effects on Lexical Complexity

Increasing task complexity had no significant effect on D value. AWU, however, increased significantly in the complex task performance. For the mediating role of affective factors, only avoidance behaviour (AB), one dimension of the learner’s writing anxiety had a significant negative moderate association with AWU in the complex task performance. The correlation between the other dimensions of writing anxiety and lexical complexity was not significant in either the simple or complex task performance. Additionally, the mediating role of motivation in AWU in writing performance did not reach statistical significance in either the simple or complex task.
4.10.4. Effects on Fluency

Increasing task complexity had an adverse significant effect on fluency. Regarding the mediating role of motivation and anxiety in L2 writing production in the simple and complex task performance, no significant association was found between the dimensions of motivation and anxiety and L2 writing fluency in the simple or complex task performance.

4.10.5. Effects on Content

Writing content scores increased as a function of increasing task complexity. As regards the mediating role of motivation and anxiety in L2 writing content in the simple and complex task performance, no significant correlation was found between the dimensions of anxiety and L2 writing content scores in the simple or complex task performance. Nonetheless, there was a significant positive moderate association between mastery goal and L2 writing content scores in the complex task performance.

4.10.6. Effects on Organisation

Increasing task complexity yielded a significant favourable effect for L2 writers’ organisation scores. The correlation between the dimensions of motivation and L2 writing organisation was not significant in either the simple or complex L2 writing task performance. The same results were obtained for the association between the dimensions of writing anxiety and writing organisation.
4.10.7. Effects on Text Quality

Writers’ writing text quality improved as a result of increasing task complexity in the level of reasoning and the number of elements. The mediating role of the dimensions of motivation and anxiety did not reach statistical significance in either the simple or complex L2 writing task performance.

4.10.8. Effects on Task Difficulty

The vast majority of the participating students and all the teachers found the six-project task more difficult than the three-project task. Iterative readings of the participants’ reasons for the complexity of the six-project (complex) task versus three-project (simple) task corroborated and extended prior findings with regard to the role of increasing the level of reasoning and number of elements in increasing the complexity of tasks.
Chapter 5.

Results for Study 2

5.1. Chapter Overview

The results of Study 2 that addressed research question (RQ) 5, the simultaneous effects of task complexity and pre-task planning on L2 writing syntactic and lexical complexity, accuracy, fluency, content, organisation, and overall text quality, are presented in this chapter. The chapter concludes with a summary of the findings.

To address research question 5, I designed and conducted Study 2 to investigate the simultaneous effects of pre-task planning conditions and task complexity on L2 writing production. The same measures that were used for measuring L2 learners’ writing production in Study 1 were employed in Study 2. Mixed ANOVAs were conducted to examine the main effects of planning conditions (present vs. absent) and task complexity, and interaction effects of planning and task complexity on each measure of writing production. Further follow-up Paired samples $t$ tests or Wilcoxon Signed Ranks tests for within group complexity effects across the simple and complex tasks and Independent Samples $t$ tests or Mann-Whitney $U$ tests for between groups across the planning conditions were run to explore further the effects of task complexity and planning conditions on writing production measures across four conditions. These conditions are simple-pre-task planning, complex-pre-task planning, simple-no-pre-task planning, and complex-no-pre-task planning. Mixed ANOVAs were not performed for the variables that did not satisfy the assumption of normality; only follow-up analyses were executed for these variables. Prior to
conducting the statistical analyses, the assumption of normality was evaluated via standard diagnostic tests and procedures (i.e., Kolmogorov-Smirnov and Shapiro-Wilk’s tests, the visual inspection of histograms, normal Q-Q, and box plots, and Skewness and Kurtosis). Furthermore, the assumption of homogeneity of variances was examined based on Levene’s $F$ test. In the following sections prior to reporting the main findings, the results of the diagnostic tests for each variable are reported.

5.2. Effects of Task Complexity and Planning Conditions on Subordinate Clauses per Clause (SUB; RQ5)

To test the impact of increasing task complexity along the level of reasoning and the number of elements and the pre-task planning versus the no-pre-task planning conditions on SUB as a measure of clausal complexity, first the normality of the learners’ scores was evaluated and found to be met (see Table 5.1). Therefore, initially, a decision was made to perform a mixed ANOVA. However, Levene's $F$ test was significant for scores for both the simple and complex task. As such, other appropriate tests were conducted to examine the effect of task complexity along the resource-directing dimension on SUB for the simple versus complex tasks under the pre-task and no-pre-task planning conditions (within groups analyses). For the resource-dispersing dimension, between groups analyses (the simple task under the pre-task planning condition versus the simple task under the no-pre-task planning condition and the complex task under the pre-task planning condition versus the complex task under the no-pre-task planning conditions) were performed.
Table 5.1. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for SUB

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBSNP</td>
<td>40</td>
<td>.007*</td>
<td>.95</td>
<td>.12</td>
<td>-.105</td>
<td>.22/08</td>
<td>.n.</td>
</tr>
<tr>
<td>SUBCNP</td>
<td>40</td>
<td>.107</td>
<td>.95</td>
<td>.14</td>
<td>-.62</td>
<td>.27/09</td>
<td>.n.</td>
</tr>
<tr>
<td>SUBSP</td>
<td>40</td>
<td>.200</td>
<td>.97</td>
<td>-.05</td>
<td>-.51</td>
<td>.29/06</td>
<td>.n.</td>
</tr>
<tr>
<td>SUBCP</td>
<td>40</td>
<td>.200</td>
<td>.95</td>
<td>.36</td>
<td>-.79</td>
<td>.36/07</td>
<td>.n.</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, SUBSNP = Subordination for simple task under no-pre-task planning condition, SUBCNP = Subordination for complex task under no-pre-task planning condition, SUBSP = Subordination for simple task under pre-task-planning condition, SUBCP = Subordination for complex task under pre-task-planning condition

The assumption of normally distributed difference scores for Paired Samples t test was met (see Table 5.2). Therefore, two Paired Samples t tests were conducted to test the differences in learners’ SUB production on the simple versus complex tasks under the pre-task and no-pre-task planning conditions (see Table 5.3).

Table 5.2. Normality Test Results for Paired Samples t Tests for SUB

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Stat.</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBNP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.4</td>
<td>.06</td>
<td>-.15</td>
<td>.73</td>
</tr>
<tr>
<td>SUBP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.1</td>
<td>.27</td>
<td>-.84</td>
<td>.73</td>
</tr>
</tbody>
</table>

Note. (S-C) = (the score on the simple task – the score on the complex task), SUBNP = subordinate clauses per clause under no-pre-task planning condition, SUBP = subordinate clauses per clause under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The learners produced significantly more SUB in the complex ($M = .36, SD = .07$) versus the simple version of the task ($M = .29, SD = .06$) under the pre-task planning condition, $t = 5.12, p < .001, d = .9$. The same was true for the learners’ SUB production on the simple ($M = .22, SD = .08$) versus the complex version of the task ($M = .27, SD = .09$) under the no-pre-task planning condition, $t = 2.76, p = .009, d = .48$. These results suggest that increasing task complexity along the level of reasoning and the number of elements enhances learners’ clausal complexity in L2 writing production; the effect for the under the pre-task planning condition was associated with a large effect size.
Table 5.3. Paired Samples t Tests Results for SUB

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUB-SP</td>
<td>40</td>
<td>.29</td>
<td>.06</td>
<td>5.12</td>
<td>39</td>
<td>.000</td>
<td>.9</td>
</tr>
<tr>
<td>SUB-CP</td>
<td>40</td>
<td>.36</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB-SNP</td>
<td>40</td>
<td>.22</td>
<td>.08</td>
<td>2.76</td>
<td>39</td>
<td>.009</td>
<td>.48</td>
</tr>
<tr>
<td>SUB-CNP</td>
<td>40</td>
<td>.27</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB-SP</td>
<td>80</td>
<td>.29</td>
<td>.06</td>
<td>4.19</td>
<td>78</td>
<td>.000</td>
<td>.79</td>
</tr>
<tr>
<td>SUB-SNP</td>
<td>80</td>
<td>.22</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB-CP</td>
<td>80</td>
<td>.36</td>
<td>.07</td>
<td>5.20</td>
<td>78</td>
<td>.000</td>
<td>.95</td>
</tr>
<tr>
<td>SUB-CNP</td>
<td>80</td>
<td>.27</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, d = effect size.

To test the differences in learners’ SUB production in the simple task under the pre-task planning ($M = .29, SD = .06$) versus the no-pre-task planning conditions ($M = .22, SD = .08$) and also in the complex version of the task under the pre-task planning ($M = .36, SD = .07$) versus the no-pre-task planning conditions ($M = .27, SD = .09$), the normality of assumption for Independent Samples $t$ test was evaluated and found to be satisfied (see Table 5.4). Hence, two Independent Samples $t$ tests were run to test the effects of providing pre-task planning on the learners’ SUB production in the simple and complex versions of the task (see Table 5.3).

Table 5.4. Normality Test Results for Independent Samples t Tests for SUB

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sig.</td>
<td>Sig.</td>
<td>Stat.</td>
<td>Std. Er</td>
<td></td>
</tr>
<tr>
<td>SUBS</td>
<td>80</td>
<td>.08</td>
<td>.05</td>
<td>.30</td>
<td>-.66</td>
<td>.53</td>
</tr>
<tr>
<td>SUBC</td>
<td>80</td>
<td>.09</td>
<td>.01</td>
<td>-.37</td>
<td>-.64</td>
<td>.53</td>
</tr>
</tbody>
</table>

Note. SUBS = Subordination for simple task under pre-task-planning and no-pre-task planning conditions, SUBC = Subordination for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal.

The results show that providing pre-task planning improved the learners’ SUB production both in the simple, $t = 4.19, p < .001, d = .79$, and in the complex versions of the task, $t = 5.20, p < .001, d = .95$. The results were associated with large effect sizes.
A graphical representation of the means and 95% confidence intervals is presented in Figure 5.1. Learners produced significantly more SUB as a function of both increasing task complexity along the level of reasoning and the number of elements, and providing pre-task planning for the simple and complex tasks.

![Graph showing SUB means and 95% confidence intervals across four conditions.]

**Figure 5.1.** Bar chart with learners’ SUB means and 95% confidence intervals across four conditions

### 5.3. Effects of Task Complexity and Pre-task Planning Conditions on Mean Length of Clauses (MLC; RQ5)

In this section the procedure and tests used to examine the influence of increasing task complexity along the level of reasoning and the number of elements and the pre-task planning versus the no-pre-task planning conditions on MLC as a measure of
sub-clausal complexity are reported. The assumption of normality for MLC scores was not met (see Table 5.5). Hence, mixed ANOVA analysis was not performed. Instead, other statistical analyses were employed to test the impact of task complexity along the *resource-directing* dimension on MLC for the simple versus complex task under the pre-task and also under the no-pre-task planning conditions (within groups analyses) and along the *resource-dispersing* dimension (pre-task versus no-pre-task planning conditions) for between groups analyses. Learners’ MLC was compared on the simple task under the pre-task planning condition versus under the no-pre-task planning condition and also on the complex task under the pre-task planning condition versus under the no-pre-task planning conditions.

### Table 5.5. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for MLC

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MLCSNP</td>
<td>40</td>
<td>.01*</td>
<td>.01*</td>
<td>.59 .37</td>
<td>-75 .73</td>
<td>9.69/2.06</td>
<td>n.n.</td>
</tr>
<tr>
<td>MLCCNP</td>
<td>40</td>
<td>.20</td>
<td>.14</td>
<td>.62 .37</td>
<td>.60 .73</td>
<td>10.01/1.96</td>
<td>n.</td>
</tr>
<tr>
<td>MLCSNP</td>
<td>40</td>
<td>.007*</td>
<td>.01*</td>
<td>.59 .37</td>
<td>-76 .73</td>
<td>9.78/2.15</td>
<td>n.n.</td>
</tr>
<tr>
<td>MLCCNP</td>
<td>40</td>
<td>.04</td>
<td>.01</td>
<td>.5 .37</td>
<td>.87 .73</td>
<td>10.17/1.92</td>
<td>n.n.</td>
</tr>
</tbody>
</table>

*Note.* Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, MLCSNP = Mean length of clauses for simple task under no-pre-task planning condition, MLCCNP = Mean length of clauses for complex task under no-pre-task planning condition, MLCSNP = Mean length of clauses for simple task under pre-task planning condition, MLCCNP = Mean length of clauses for complex task under pre-task planning condition.

As the assumption of normality was met (see Table 5.6), two Paired Samples *t* tests were carried out to check the differences in learners’ MLC on the simple versus the complex versions of the task under the pre-task and no-pre-task planning conditions (see Table 5.7).
Table 5.6. Normality Test Results for Paired Samples $t$ Tests for MLC

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MLCNP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.38</td>
<td>.03</td>
<td>.37</td>
<td>-.09</td>
<td>.73</td>
<td>n.</td>
</tr>
<tr>
<td>MLCP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.78</td>
<td>-.04</td>
<td>.37</td>
<td>-.29</td>
<td>.73</td>
<td>n.</td>
</tr>
</tbody>
</table>

*Note.* (S-C) = (the score on the simple task – the score on the complex task), MLCNP = mean length of clauses under no-pre-task planning condition, MLCP = mean length of clauses under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

Learners’ MLC did not differ significantly in the complex ($M = 10.17, SD = 1.92$) versus the simple ($M = 9.78, SD = 2.15$) version of the task under the pre-task planning condition, $t(39) = 1.30, p = .20, d = .16$. Likewise, the learners’ MLC on the simple ($M = 9.69, SD = 2.06$) versus the complex ($M = 10.01, SD = 1.96$) versions of the task under the no-pre-task planning condition did not vary significantly, $t(39) = 1.15, p = .25, d = .18$. These findings indicate that increasing task complexity along the level of reasoning and the number of elements did not lead to improvements in learners’ sub-clausal complexity in L2 writing production.

Table 5.7. Paired Samples $t$ Tests Results for MLC

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$t$</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC-SP</td>
<td>40</td>
<td>9.78</td>
<td>2.15</td>
<td>1.30</td>
<td>39</td>
<td>.20</td>
<td>.16</td>
</tr>
<tr>
<td>MLC-CP</td>
<td>40</td>
<td>10.17</td>
<td>1.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLC-SNP</td>
<td>40</td>
<td>9.69</td>
<td>2.06</td>
<td>1.15</td>
<td>39</td>
<td>.25</td>
<td>.18</td>
</tr>
<tr>
<td>MLC-CNP</td>
<td>40</td>
<td>10.01</td>
<td>1.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, $d$ = effect size

The differences in learners’ MLC in the simple task under the pre-task planning ($M = 9.78, SD = 2.15$) versus the no-pre-task planning conditions ($M = 9.69, SD = 2.06$) and also in the complex version of the task under the pre-task planning ($M = 10.17, SD = 1.92$) versus the no-pre-task planning conditions ($M = 10.01, SD = 1.96$) were
examined via two Mann-Whitney U tests as the normality assumption was not met (see Table 5.8).

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Std. Er</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLCS</td>
<td>80</td>
<td>.06*</td>
<td>.005*</td>
<td>.60</td>
<td>.26</td>
<td>-.78</td>
</tr>
<tr>
<td>MLCC</td>
<td>80</td>
<td>.000</td>
<td>.000*</td>
<td>.53</td>
<td>.26</td>
<td>.41</td>
</tr>
</tbody>
</table>

*Note. MLCS = mean length of clauses for simple task under pre-task-planning and no-pre-task planning conditions, MLCC = mean length of clauses for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The results indicate providing pre-task planning did not improve the learners’ MLC either in the simple, \( z = .19, p = .43, r = .02 \), or in the complex versions of the task, \( z = .84, p = .66, r = .11 \), (see Table 5.9).

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>z</th>
<th>Df</th>
<th>Sig.(2-tailed)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC-SP</td>
<td>40</td>
<td>9.78</td>
<td>2.15</td>
<td>.19</td>
<td>78</td>
<td>.43</td>
<td>.02</td>
</tr>
<tr>
<td>MLC-SNP</td>
<td>40</td>
<td>9.69</td>
<td>2.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLC-CP</td>
<td>40</td>
<td>10.16</td>
<td>1.92</td>
<td>.84</td>
<td>78</td>
<td>.66</td>
<td>.11</td>
</tr>
<tr>
<td>MLC-CNP</td>
<td>40</td>
<td>10.01</td>
<td>1.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, r = effect size

A graphical representation of the means and 95% confidence intervals is presented in Figure 5.2. Neither increasing task complexity along the level of reasoning and the number of elements nor providing pre-task planning for the simple and complex tasks augmented learners’ sub-clausal complexity in L2 writing production.
5.4. Effects of Task Complexity and Pre-task Planning Conditions on Phrasal Coordination (COR; RQ5)

The results of the assumption of normality tests revealed that mixed ANOVA might not be an appropriate test for COR scores (see Table 5.10). Therefore, statistical analyses were run to test the effect of task complexity along the resource-directing dimension (within groups analyses for the simple versus complex tasks under the pre-task and also under the no-pre-task planning conditions) and along the resource-dispersing dimension (between groups analyses for pre-task versus no-pre-task planning conditions) on COR.

Figure 5.2. Bar chart with learners’ MLC means and 95% confidence intervals across four conditions
Table 5.10. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for COR

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>M/SD</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORSNP</td>
<td>40</td>
<td>.03*</td>
<td>.001*</td>
<td>1.25</td>
<td>1.59</td>
<td>.73</td>
<td>.28/.17</td>
</tr>
<tr>
<td>CORCNP</td>
<td>40</td>
<td>.000*</td>
<td>.000*</td>
<td>2.36</td>
<td>7.58</td>
<td>.73</td>
<td>.29/22</td>
</tr>
<tr>
<td>CORSP</td>
<td>40</td>
<td>.02*</td>
<td>.001*</td>
<td>1.14</td>
<td>1.20</td>
<td>.73</td>
<td>.27/17</td>
</tr>
<tr>
<td>CORCP</td>
<td>40</td>
<td>.000*</td>
<td>.000*</td>
<td>1.22</td>
<td>1.09</td>
<td>.73</td>
<td>.27/18</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, CORSNP = Coordination for simple task under no-pre-task planning condition, CORCNP = Coordination for complex task under no-pre-task planning condition, CORSP = Coordination for simple task under pre-task-planning condition, CORCP = Coordination for complex task under pre-task-planning condition

The assumption of normally distributed difference COR scores for Paired Samples t test was not met (see Table 5.11). Therefore, two Wilcoxon Signed Rank tests were used for within groups analyses.

Table 5.11. Normality Test Results for Paired Samples t Tests for COR

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORNP(S-C)</td>
<td>40</td>
<td>.001*</td>
<td>.000</td>
<td>-1.98</td>
<td>8.86</td>
<td>.73</td>
</tr>
<tr>
<td>CORP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.08</td>
<td>.73</td>
<td>1.31</td>
<td>.73</td>
</tr>
</tbody>
</table>

Note. (S-C) = (the score on the simple task – the score on the complex task), CORNP = phrasal coordination per t-unit under no-pre-task planning condition, CORP = phrasal coordination per t-unit under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The results revealed that learners’ did not produce significantly more COR in the complex ($M = .27$, $SD = .18$) relative to the simple version of the task ($M = .27$, $SD = .17$) under the pre-task planning condition, $z = -.50$, $p = .62$, $r = .05$. The same trend was found for the learners’ COR production in the simple ($M = .28$, $SD = .17$) versus the complex version of the task ($M = .29$, $SD = .22$) under the no-pre-task planning condition, $z = -.75$, $p = .45$, $r = .08$ (see Table 5.12). These results indicate that increasing task complexity along the level of reasoning and the number of elements did not contribute to improving phrasal coordination in L2 writing production.
Table 5.12. Wilcoxon Signed Rank Test and Mann-Whitney U Tests Results for COR

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>z</th>
<th>Df</th>
<th>Sig. (2-tailed)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>COR-SP</td>
<td>40</td>
<td>.27</td>
<td>.17</td>
<td>-.50</td>
<td>39</td>
<td>.62</td>
<td>.05</td>
</tr>
<tr>
<td>COR-CP</td>
<td>40</td>
<td>.27</td>
<td>.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COR-SNP</td>
<td>40</td>
<td>.28</td>
<td>.17</td>
<td>-.75</td>
<td>39</td>
<td>.45</td>
<td>.08</td>
</tr>
<tr>
<td>COR-CNP</td>
<td>40</td>
<td>.29</td>
<td>.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COR-SP</td>
<td>40</td>
<td>.27</td>
<td>.17</td>
<td>-.27</td>
<td>78</td>
<td>.78</td>
<td>.03</td>
</tr>
<tr>
<td>COR-SNP</td>
<td>40</td>
<td>.28</td>
<td>.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COR-CP</td>
<td>40</td>
<td>.27</td>
<td>.18</td>
<td>-.20</td>
<td>78</td>
<td>.84</td>
<td>.02</td>
</tr>
<tr>
<td>COR-CNP</td>
<td>40</td>
<td>.29</td>
<td>.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, r = effect size

To test the differences in learners’ COR production in the simple task under the pre-task planning ($M = .27, SD = .17$) versus the no-pre-task planning condition ($M = .28, SD = .17$) and also in the complex version of the task under the pre-task planning ($M = .27, SD = .18$) versus the no-pre-task planning condition ($M = .29, SD = .22$), two Mann-Whitney U tests were performed. The assumption of normality for Independent Samples $t$ tests was not met (see Table 5.13).

Table 5.13. Normality Test Results for Independent Samples $t$ Tests for COR

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Std. Er.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORS</td>
<td>80</td>
<td>.000*</td>
<td>.000*</td>
<td>1.17</td>
<td>1.25</td>
</tr>
<tr>
<td>CORC</td>
<td>80</td>
<td>.000*</td>
<td>.000*</td>
<td>1.97</td>
<td>5.66</td>
</tr>
</tbody>
</table>

Note. CORS = Coordination for simple task under pre-task-planning and no-pre-task planning conditions, CORC = Coordination for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The results show that providing pre-task planning did not have a beneficial effect on the learners’ COR production in either the simple, $z = -.27, p = .78, r = .03$ or in the complex version of the task, $z = -.20, p = .84, r = .02$ (see Table 5.16).

A graphical illustration of the means and 95% confidence intervals is presented in Figure 5.3. The learners did not use significantly more COR in their L2 writing
production as a function of either increasing task complexity along the level of reasoning and the number of elements, or providing pre-task planning for the simple and complex tasks.

![Bar chart with learners' COR means and 95% confidence intervals across four conditions](image)

**Figure 5.3.** Bar chart with learners’ COR means and 95% confidence intervals across four conditions

### 5.5. Effects of Task Complexity and Planning Conditions on the Number of Errors per T-unit (EPTU; RQ5)

In order to test the effects of the pre-task planning conditions and increasing task complexity in terms of the degree of reasoning and the number of elements on English-as-a-foreign-language accuracy (as measured by the number of errors per t-units), a mixed ANOVA and follow-up Paired Samples *t* tests and Independent Samples *t* tests were run. The assumption of normality was met (see Table 5.14) and Levene’s *F* test for EPTUS (the number of errors per t-unit for the simple task), *F*(1, 78) = .03, *p* = .95, and for EPTUC (the number of errors per t-unit for the complex task), *F*(1, 78) = .09, *p* = .77, was not significant.
Table 5.14. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for EPTU

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>M/SD</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPTUSNP</td>
<td>40</td>
<td>.20</td>
<td>.08</td>
<td>.70</td>
<td>.37</td>
<td>.06/.73</td>
<td>1.27/.52</td>
</tr>
<tr>
<td>EPTUCNP</td>
<td>40</td>
<td>.20</td>
<td>.93</td>
<td>.15</td>
<td>.37</td>
<td>.17/.73</td>
<td>1.71/.70</td>
</tr>
<tr>
<td>EPTUSP</td>
<td>40</td>
<td>.20</td>
<td>.29</td>
<td>.50</td>
<td>.37</td>
<td>-.34/.73</td>
<td>1.31/.50</td>
</tr>
<tr>
<td>EPTUCP</td>
<td>40</td>
<td>.20</td>
<td>.69</td>
<td>.40</td>
<td>.37</td>
<td>.08/.73</td>
<td>1.75/.66</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er.= Error, n. = Normal, n.n. = Non-normal, EPTUSNP = Errors per t-unit for simple task under no-pre-task planning condition, EPTUCNP = Errors per t-unit for complex task under no-pre-task planning condition, EPTUSP = Errors per t-unit for simple task under pre-task-planning condition, EPTUCP = Errors per t-unit for complex task under pre-task-planning condition

The descriptive statistics associated with the learners’ EPTU across four conditions are reported in Table 5.14. As can be seen, complex-pre-task planning, complex no-pre-task-planning, simple-pre-task planning, and simple-no-pre-task planning conditions were associated with the numerically highest to lowest mean level of learners’ EPTU ($M = 1.75$, $SD = .66$), ($M = 1.71$, $SD = .70$), ($M = 1.31$, $SD = .50$), and ($M = 1.27$, $SD = .52$), respectively.

The mixed ANOVA yielded a significant main effect for increasing task complexity along the resource-directing dimension (increasing the level of reasoning and the number of elements), $F(1, 78) = 44.59$, $p < .001$, partial eta-squared ($\eta^2_p$) = .37. However, the learners’ production under the pre-task versus no-pre-tasking planning conditions, $F(1, 78) = .09$, $p = .89$, partial eta-squared ($\eta^2_p$) = .002, and the interaction effect, $F(1, 78) = .05$, $p = .85$, partial eta-squared ($\eta^2_p$) = .001, did not reach statistical significance. These findings indicate that increasing task complexity along the resource-directing dimension (collapsing across the planning conditions) increased the number of errors the learners made per t-unit, and providing pre-task-planning (collapsing across the resource-directing dimension) did not affect their accuracy.
significantly. Additionally, the learners’ level of accuracy as measured via EPTU on the simple versus the complex versions of the task was not dependent on planning conditions.

To test the differences between the learners’ production on the simple \((M = 1.27, SD = .52)\) versus the complex version of the task \((M = 1.71, SD = .70)\) under the no-pre-task-planning condition, and on the same simple \((M = 1.31, SD = .50)\) versus the same complex task \((M = 1.75, SD = .66)\) under pre-task-planning condition, two Paired Samples \(t\) tests were performed. The assumption of normally distributed difference scores was met (see Table 5.15).

**Table 5.15. Normality Test Results for Paired Samples \(t\) Tests for EPTU**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ERPTNP(S-C)</td>
<td>40</td>
<td>.05</td>
<td>.31</td>
<td>.25</td>
<td>.37</td>
<td>.62</td>
<td>.73</td>
<td>n</td>
</tr>
<tr>
<td>ERPTP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.36</td>
<td>-.17</td>
<td>.37</td>
<td>.03</td>
<td>.73</td>
<td>n</td>
</tr>
</tbody>
</table>

Note. (S-C) = (the score on the simple task – the score on the complex task), ERPTNP = Errors per t-unit under no-pre-task planning condition, ERPTP = Errors per t-unit under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The Paired Samples \(t\) tests and Independent Samples \(t\) tests results are reported in Table 5.17. Under both the no-pre-task planning condition, \(t(39) = -4.52, p < .001, d = .70\), and the pre-task planning condition, \(t(39) = -4.93, p < .001, d = .78\), increasing task complexity had detrimental effects on the learners’ accuracy. The results for both conditions were significant and associated with relatively large effect sizes.

To test the differences between the learners’ accuracy in terms of ERPT on the simple version of the task under the pre-task planning condition \((M = 1.31, SD = .50)\) versus the no-pre-task planning condition \((M = 1.27, SD = .52)\) and also the
difference in learners’ ERPT on the complex version of the task under the no-pre-task planning condition ($M = 1.71, SD = .70$) and on the same task under the pre-task planning condition ($M = 1.75, SD = .66$), two Independent Samples $t$ tests were executed. The assumption of normality was met (see Table 5.16).

Table 5.16. Normality Test Results for Independent Samples $t$ Tests for ERPT

<table>
<thead>
<tr>
<th>Variables</th>
<th>$N$</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sig.</td>
<td>Sig.</td>
<td>Stat.</td>
<td>Std. Er.</td>
</tr>
<tr>
<td>ERPTS</td>
<td>80</td>
<td>.20</td>
<td>.01*</td>
<td>.60</td>
<td>.27</td>
</tr>
<tr>
<td>ERPTC</td>
<td>80</td>
<td>.08</td>
<td>.40</td>
<td>.25</td>
<td>.26</td>
</tr>
</tbody>
</table>

Note. ERPTS = Errors per t-unit for simple task under pre-task-planning and no-pre-task planning conditions, ERPTC = Errors per t-unit for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n., = Non-normal

As displayed in Table 5.17, providing pre-task planning did not yield a significant difference in learners’ ERPT on the simple, $t(78) = .32, p = .74, d = .07$, versus the complex task, $t(78) = .26, p = .79, d = .06$.

Table 5.17. Paired and Independent Samples $t$ Tests Results for ERPT

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$T$</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERPT-SP</td>
<td>40</td>
<td>1.31</td>
<td>.50</td>
<td>-4.93</td>
<td>39</td>
<td>.000</td>
<td>.78</td>
</tr>
<tr>
<td>ERPT-CP</td>
<td>40</td>
<td>1.75</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERPT-SNP</td>
<td>40</td>
<td>1.27</td>
<td>.52</td>
<td>-4.52</td>
<td>39</td>
<td>.000</td>
<td>.70</td>
</tr>
<tr>
<td>ERPT-CNP</td>
<td>40</td>
<td>1.71</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERPT-SP</td>
<td>40</td>
<td>1.31</td>
<td>.50</td>
<td>.32</td>
<td>78</td>
<td>.74</td>
<td>.07</td>
</tr>
<tr>
<td>ERPT-SNP</td>
<td>40</td>
<td>1.27</td>
<td>.52</td>
<td>.26</td>
<td>78</td>
<td>.79</td>
<td>.06</td>
</tr>
<tr>
<td>ERPT-CNP</td>
<td>40</td>
<td>1.71</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, $d$ = effect size

A visual representation of the means and 95% confidence intervals is presented in Figure 5.4. As displayed in the Figure, learners’ accuracy deteriorated as a function of increasing task complexity along the reasoning and the number of elements under both the pre-task planning ($d = .78$) and no-pre-task planning conditions ($d = .70$). The effect sizes for both conditions were reasonably large. However, the provision of
pre-task-planning did not lead to improvements in learners’ level of accuracy in L2 writing production.

Figure 5.4. Bar chart with ERPT means and 95% confidence intervals across four conditions

5.6. Effects of Task Complexity and Planning Conditions on Error-free T-units per T-unit (EFTPT; RQ5)

In order to test the effects of pre-task planning conditions and increasing task complexity regarding the level of reasoning and the number of elements on English-as-a-foreign-language accuracy as measured by the number of error-free t-units per t-unit, which was employed to capture the task complexity effect on the absence of errors, a mixed ANOVA and follow-up Paired Samples $t$ tests and Independent Samples $t$ tests were run. The assumptions of normality was met (see Table 5.18) and Levene’s $F$ test for EFTPTS (error-free t-units per t-unit for the simple task), $F(1, 78) = .04, p = .89$, and for EFTPTC (error-free t-units per t-unit for the complex task), $F(1, 78) = .05, p = .82$, was not significant.
Table 5.18. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for EFTPT

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EFTPTSNP</td>
<td>40</td>
<td>.20</td>
<td>.35</td>
<td>-.14</td>
<td>.37</td>
<td>-.30</td>
<td>.73</td>
<td>.31/.16</td>
<td>n</td>
</tr>
<tr>
<td>EFTPTCNP</td>
<td>40</td>
<td>.20</td>
<td>.14</td>
<td>.65</td>
<td>.37</td>
<td>1.10</td>
<td>.73</td>
<td>.27/.14</td>
<td>n</td>
</tr>
<tr>
<td>EFTPTSP</td>
<td>40</td>
<td>.20</td>
<td>.36</td>
<td>-.16</td>
<td>.37</td>
<td>-.41</td>
<td>.73</td>
<td>.32/17</td>
<td>n</td>
</tr>
<tr>
<td>EFTPTCP</td>
<td>40</td>
<td>.20</td>
<td>.20</td>
<td>.68</td>
<td>.37</td>
<td>1.08</td>
<td>.73</td>
<td>.26/.13</td>
<td>n</td>
</tr>
</tbody>
</table>

*Note.* Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, EFTPTSNP = Errors free t-units per t-unit for simple task under no-pre-task planning condition, EFTPTCNP = Errors free t-units per t-unit for complex task under no-pre-task planning condition, EFTPTSP = Errors free t-units per t-unit for simple task under pre-task-planning condition, EFTPTCP = Errors free t-units per t-unit for complex task under pre-task-planning condition

The descriptive statistics associated with the learners’ EFTPT across four conditions are reported in Table 5.18. The Mixed ANOVA yielded a significant main effect for increasing task complexity along the *resource-directing* dimension (increasing the level of reasoning and the number of elements), *F*(1, 78) = 6.76, *p* < .001, partial eta-squared (*η²*) = .08. For the pre-task versus no-pre-tasking planning conditions, *F*(1, 78) = .04, *p* = .84, partial eta-squared (*η²*) = .001, and for the interaction effect, *F*(1, 78) = .05, *p* = .82, partial eta-squared (*η²*) = .001, the results were not significant.

These findings suggest that increasing task complexity along the *resource-directing* dimension (collapsing across the planning conditions) reduced learners’ level of accuracy in terms of the number of error-free t-units per t-unit significantly in learners’ L2 writing production. Nevertheless, the availability of pre-task-planning (collapsing across the resource-directing dimension) did not influence learners’ accuracy. Thus, the learners’ level of accuracy in the two versions (simple versus complex) of the task was not contingent upon planning conditions.
As the assumption of normality was met, two Paired Samples \( t \) tests were carried out to check the difference in learners’ EFTPT on the simple versus the complex versions of the task under the pre-task and no-pre-task planning conditions (see Table 5.19).

**Table 5.19. Normality Test Results for Paired Samples \( t \) Tests for EFTPT**

<table>
<thead>
<tr>
<th>Variables</th>
<th>( N )</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFTPTNP(S-C)</td>
<td>40</td>
<td>.20</td>
<td>.16</td>
<td>-.73</td>
<td>.37</td>
<td>.54 .73 n</td>
</tr>
<tr>
<td>EFTPTP(S-C)</td>
<td>40</td>
<td>.14</td>
<td>.23</td>
<td>-.68</td>
<td>.37</td>
<td>.64 .73 n</td>
</tr>
</tbody>
</table>

*Note.* (S-C) = (the score on the simple task – the score on the complex task), EFTPTNP = Error-free \( t \)-units per \( t \)-unit under no-pre-task planning condition, EFTPTP = Error-free \( t \)-units per \( t \)-unit under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The learners produced significantly more EFTPT in the simple (\( M = .32, SD = .17 \)) compared to the complex version of the task (\( M = .26, SD = .13 \)) under the pre-task planning condition, \( t(39) = 2.04, p < .05, d = .29 \). However, the learner’s EFTPT production on the simple (\( M = .31, SD = .16 \)) versus the complex versions of the task (\( M = .27, SD = .14 \)) did not vary significantly under the no-pre-task planning condition, \( t(39) = 1.64, p = .10, d = .26 \) (see Table 5.20). The results show learners’ accuracy deteriorated significantly only under the pre-task planning condition as a function of increasing task complexity. Nonetheless, the results for both conditions were associated with small effect sizes.

**Table 5.20. Paired Samples \( t \) Tests Results for EFTPT**

<table>
<thead>
<tr>
<th></th>
<th>( N )</th>
<th>( M )</th>
<th>( SD )</th>
<th>( T )</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFTPT-SP</td>
<td>40</td>
<td>.32</td>
<td>.17</td>
<td>2.04</td>
<td>39</td>
<td>.04</td>
<td>.29</td>
</tr>
<tr>
<td>EFTPT-CP</td>
<td>40</td>
<td>.26</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EFTPT-SNP</td>
<td>40</td>
<td>.31</td>
<td>.16</td>
<td>1.64</td>
<td>39</td>
<td>.10</td>
<td>.26</td>
</tr>
<tr>
<td>EFTPT-CNP</td>
<td>40</td>
<td>.27</td>
<td>.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, \( d \) = effect size
For between groups analyses, the differences in learners’ EFTPT in the simple task under the pre-task planning \((M = .32, SD = .17)\) versus no-pre-task planning conditions \((M = .31, SD = .16)\) and also in the complex version of the task under the pre-task planning \((M = .26, SD = .13)\) versus no-pre-task planning conditions \((M = .27, SD = .14)\) were evaluated. As the normality assumption was violated (see Table 5.21), two Mann-Whitney \(U\) tests were carried out to test the effects of providing pre-task planning on the learners’ EFTPT production in the simple and the complex versions of the task.

**Table 5.21.** Normality Test Results for Independent Samples \(t\) Tests for EFTPT

<table>
<thead>
<tr>
<th>Variables</th>
<th>(N)</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFTPTS</td>
<td>80</td>
<td>.20</td>
<td>.06</td>
<td>-.15</td>
<td>.26</td>
<td>.41</td>
</tr>
<tr>
<td>EFTPTC</td>
<td>80</td>
<td>.20</td>
<td>.01*</td>
<td>.66</td>
<td>.26</td>
<td>.96</td>
</tr>
</tbody>
</table>

*Note. EFTPTS = Error-free t-units per t-unit for simple task under pre-task-planning and no-pre-task planning conditions, EFTPTC = Error-free t-units per t-unit for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal*

As can be seen from Table 5.22, these results indicate that providing pre-task planning did not improve the learners’ EFTPT in the simple, \(z = -.09, p = .92, r = .01\), or complex version of the task, \(z = -.35, p = .72, r = .04\).

**Table 5.22.** Mann-Whitney \(U\) Tests Results for EFTPT

<table>
<thead>
<tr>
<th></th>
<th>(N)</th>
<th>(M)</th>
<th>(SD)</th>
<th>(Z)</th>
<th>(df)</th>
<th>Sig.(2-tailed)</th>
<th>(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFTPT-SP</td>
<td>40</td>
<td>.32</td>
<td>.17</td>
<td>-.09</td>
<td>78</td>
<td>.92</td>
<td>.01</td>
</tr>
<tr>
<td>EFTPT-SNP</td>
<td>40</td>
<td>.31</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EFTPT-CP</td>
<td>40</td>
<td>.26</td>
<td>.13</td>
<td>-.35</td>
<td>78</td>
<td>.72</td>
<td>.04</td>
</tr>
<tr>
<td>EFTPT-CNP</td>
<td>40</td>
<td>.27</td>
<td>.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, \(r\) = effect size*

A graphical illustration of the means and 95% confidence intervals is presented in Figure 5.5. The learners’ accuracy decreased as a function of increasing task
complexity along the *resource directing* dimension (ignoring pre-task planning conditions); however, provision of pre-task planning did not affect L2 writing accuracy significantly.

![Figure 5.5](image.png)

**Figure 5.5.** Bar chart with EFTPT means and 95% confidence intervals across four conditions

### 5.7. Effects of Task Complexity and Planning Conditions on Lexical Diversity (LD; RQ5)

A mixed ANOVA and follow-up analyses were performed to test the effects of pre-task planning conditions and increasing task complexity along the level of reasoning and the number of elements on English-as-a-foreign-language (EFL) learners’ lexical diversity (LD) in L2 writing production. A mixed ANOVA was run as the assumption of normality was met (see Table 5.23) and Levene’s *F* test for LDS
(lexical diversity for the simple task), $F(1, 78) = .003$, $p = .95$, and LDC (lexical diversity for the complex task), $F(1, 78) = .015$, $p = .90$, was not significant.

Table 5.23. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for LD

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LDSNP</td>
<td>40</td>
<td>.20</td>
<td>.58</td>
<td>-24 .37</td>
<td>-.18 .73</td>
<td>75.70/12.29</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDCNP</td>
<td>40</td>
<td>.20</td>
<td>.79</td>
<td>-10 .37</td>
<td>.08 .73</td>
<td>77.10/11.31</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDSP</td>
<td>40</td>
<td>.20</td>
<td>.39</td>
<td>-29 .37</td>
<td>-.24 .73</td>
<td>75.94/11.98</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDCP</td>
<td>40</td>
<td>.20</td>
<td>.75</td>
<td>-13 .37</td>
<td>-.03 .73</td>
<td>76.82/10.94</td>
<td>n</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er.= Error, n. = Normal, n.n. = Non-normal, LDSNP = Lexical diversity for simple task under no-pre-task planning condition, LDCNP = Lexical diversity for complex task under no-pre-task planning condition, LDSP = Lexical diversity for simple task under pre-task-planning condition, LDCP = Lexical diversity for complex task under pre-task-planning condition

The descriptive statistics for the learners’ LD across four conditions are reported in Table 5.23. The ordering of mean LD scores from highest to lowest was: Complex no-pre-task-planning ($M = 77.10$, $SD = 11.31$), complex-pre-task planning ($M = 76.82$, $SD = 10.94$), simple-pre-task planning ($M = 75.94$, $SD = 11.98$) and simple-no-pre-task planning conditions ($M = 75.70$, $SD = 12.29$), respectively.

The mixed ANOVA results for main effects for increasing task complexity along the resource-directing dimension (increasing the level of reasoning and the number of elements), $F(1, 78) = 3.68$, $p = .06$, partial eta-squared ($\eta^2_p$) = .04, for the learners’ LD under the pre-task versus no-pre-tasking planning conditions, $F(1, 78) = .15$, $p = .72$, partial ($\eta^2_p$) = .001, and for the interaction effect, $F(1, 78) = .19$, $p = .67$, partial eta-squared ($\eta^2_p$) = .002, were not significant. These results show that neither increasing task complexity along the resource-directing dimension (collapsing across the planning conditions) nor providing pre-task-planning (collapsing across the resource-directing dimension) enhanced LD in L2 writing production significantly.
Furthermore, LD in L2 writing production tasks was not contingent upon the availability of planning time.

To evaluate the nature of differences across four conditions, the mixed ANOVA analysis was followed up by Wilcoxon Signed Rank tests for within-groups and Independent Samples t tests for between-groups conditions. To test the differences between the learners’ production on the simple versus the complex version of the task under the no-pre-task-planning condition, and on the same simple versus the same complex task under the pre-task-planning condition, two Wilcoxon Signed Rank tests were performed. The assumption of normally distributed difference scores was not met (see Table 5.24).

**Table 5.24. Normality Test Results for Paired Samples t Tests for LD**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LDNP(S-C)</td>
<td>40</td>
<td>.01</td>
<td>.06</td>
<td>.75</td>
<td>.37</td>
<td>.68</td>
<td>.73</td>
<td>n.n.</td>
</tr>
<tr>
<td>LDP(S-C)</td>
<td>40</td>
<td>.02</td>
<td>.07</td>
<td>.69</td>
<td>.37</td>
<td>1.53</td>
<td>.73</td>
<td>n.n.</td>
</tr>
</tbody>
</table>

*Note.* (S-C) = (the score on the simple task – the score on the complex task), LDNP = Lexical diversity under no-pre-task planning condition, LDP = Lexical diversity under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The results were not significant either in the complex ($M = 76.82$, $SD = 10.94$) compared to the simple version of the task ($M = 75.94$, $SD = 11.98$) under the pre-task planning condition, $z = -1.71$, $p = .09$, $r = .19$, or on the simple ($M = 75.70$, $SD = 12.29$) versus the complex version of the task ($M = 77.10$, $SD = 11.13$) under the no-pre-task planning condition, $z = -1.93$, $p = .05$, $r = .21$. This suggests that neither the availability of planning time nor increasing task complexity improved LD in L2 writing production.
Table 5.25. Wilcoxon Signed Rank Test Results for LD

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Z</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD-SP</td>
<td>40</td>
<td>75.94</td>
<td>11.98</td>
<td>-1.71</td>
<td>39</td>
<td>.09</td>
<td>.19</td>
</tr>
<tr>
<td>LD-CP</td>
<td>40</td>
<td>76.82</td>
<td>10.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD-SNP</td>
<td>40</td>
<td>75.70</td>
<td>12.29</td>
<td>-1.93</td>
<td>39</td>
<td>.05</td>
<td>.21</td>
</tr>
<tr>
<td>LD-CNP</td>
<td>40</td>
<td>77.10</td>
<td>11.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, r = effect size

To test the effects of availability of the pre-task planning on LD in L2 writing production, the learners’ LD on the simple version of task under the pre-task planning condition (\(M = 75.94, SD = 11.98\)) versus the no-pre-task planning condition (\(M = 75.70, SD = 12.29\)) and on the complex version of the task under the no-pre-task planning condition (\(M = 77.10, SD = 11.13\)) versus the pre-task planning condition (\(M = 76.82, SD = 10.94\)) were analysed. Two Independent Samples \(t\) tests were run, as the assumption of normality was met (see Table 5.26).

Table 5.26. Normality Test Results for Independent Samples \(t\) Tests for LD

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Stat.</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDS</td>
<td>80</td>
<td>.20</td>
<td>.08</td>
<td>-.26</td>
<td>-.27</td>
<td>.53</td>
</tr>
<tr>
<td>LDC</td>
<td>80</td>
<td>.19</td>
<td>.26</td>
<td>-.11</td>
<td>-.05</td>
<td>.53</td>
</tr>
</tbody>
</table>

Note. LDS = Lexical diversity for simple task under pre-task-planning and no-pre-task planning conditions, LDC = Lexical diversity for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

As displayed in Table 5.27, the results show that providing pre-task planning did not increase LD in L2 writing production significantly on either the simple, \(t(78) = .09, p = .93, d = .02\), or complex tasks, \(t(78) = -.11, p = .90, d = .03\).
Table 5.27. Independent Samples t Tests Results for LD

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>T</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD-SP</td>
<td>40</td>
<td>75.94</td>
<td>11.98</td>
<td>.09</td>
<td>78</td>
<td>.93</td>
<td>.02</td>
</tr>
<tr>
<td>LD-SNP</td>
<td>40</td>
<td>75.70</td>
<td>12.69</td>
<td>-.11</td>
<td>78</td>
<td>.90</td>
<td>.03</td>
</tr>
<tr>
<td>LD-CP</td>
<td>40</td>
<td>76.82</td>
<td>10.94</td>
<td>-.11</td>
<td>78</td>
<td>.90</td>
<td>.03</td>
</tr>
<tr>
<td>LD-CNP</td>
<td>40</td>
<td>77.10</td>
<td>11.13</td>
<td>-.11</td>
<td>78</td>
<td>.90</td>
<td>.03</td>
</tr>
</tbody>
</table>

*Note.* SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, \(d\) = effect size

A visual depiction of the means and 95% confidence intervals is presented in Figure 5.6. As displayed in the Figure, neither increasing task complexity along the reasoning and the number of elements nor provision of pre-task planning enhanced LD in L2 writing production.

![Bar chart with learners’ LD means and 95% confidence intervals across four conditions](image)

**Figure 5.6.** Bar chart with learners’ LD means and 95% confidence intervals across four conditions
5.8. Effects of Task Complexity and Planning Conditions on Academic Word Use (AWU; RQ5)

The assumption of normality for AWU scores was not met (see Table 5.28). Hence, a mixed ANOVA was not performed, and analyses were performed to test the effects of task complexity along the resource-directing dimension on AWU for the simple versus the complex task under the pre-task and the no-pre-task planning conditions (within groups analysis) and along the resource-dispersing dimension (the simple task under the pre-task planning condition versus the no-pre-task planning condition and the complex task under the pre-task planning condition versus the no-pre-task planning condition).

Table 5.28. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for AWU

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AWUSNP</td>
<td>40</td>
<td>.002*</td>
<td>.001*</td>
<td>.892 .37</td>
<td>-.145 .73</td>
<td>.96/.90 n.n.</td>
</tr>
<tr>
<td>AWUCNP</td>
<td>40</td>
<td>.041*</td>
<td>.037*</td>
<td>.63 .37</td>
<td>-.39 .73</td>
<td>1.57/.74 n.</td>
</tr>
<tr>
<td>AWUSP</td>
<td>40</td>
<td>.002*</td>
<td>.000*</td>
<td>1.102 .37</td>
<td>.29 .73</td>
<td>.84/.84 n.n.</td>
</tr>
<tr>
<td>AWUCP</td>
<td>40</td>
<td>.171*</td>
<td>.303*</td>
<td>.471 .37</td>
<td>-.02 .73</td>
<td>1.70/.85 n.n.</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er = Error, n. = Normal, n.n. = Non-normal, AWUSNP = Academic vocabulary use for simple task under no-pre-task planning condition, AWUCNP = Academic vocabulary use for complex task under no-pre-task planning condition, AWUSP = Academic vocabulary use for simple task under pre-task-planning condition, AWUCP = Academic vocabulary use for complex task under pre-task-planning condition

The assumption of normally distributed difference scores for Paired Samples t test was not met (see Table 5.29). Therefore, two Wilcoxon Signed Rank tests were conducted to test the differences in learners’ AWU on the simple versus the complex task under the pre-task and the no-pre-task planning conditions.
Learners’ AWU increased significantly in the complex task ($M = 1.70, SD = .85$) compared to the simple version of the task ($M = .84, SD = .84$) under the pre-task planning condition, $z = -4.9, p < .001, r = .55$. The same trend was found for the learners’ AWU on the simple ($M = .96, SD = .90$) versus the complex version of the task ($M = 1.57, SD = .74$) under the no-pre-task planning condition, $z = -4.1, p < .001, r = .46$ (see Table 5.30). These results indicate the beneficial role of increasing task complexity along the level of reasoning and the number of elements in learners’ AWU in L2 writing production.

To address the potential differences in learners’ AWU in the simple task under the pre-task planning ($M = .84, SD = .84$) versus the no-pre-task planning condition ($M = .96, SD = .90$) and also in the complex version of the task under the pre-task planning ($M = 1.70, SD = .85$) versus the no-pre-task planning conditions ($M = 1.57, SD = .74$),
.74), two Mann-Whitney U tests were performed. The normality of assumption for Independent Samples t Test was violated (see Table 5.31).

**Table 5.31. Normality Test Results for Independent Samples t Tests for AWU**

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Stat.</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWUS</td>
<td>80</td>
<td>.000*</td>
<td>.000</td>
<td>.96</td>
<td>-.02</td>
<td>.53</td>
</tr>
<tr>
<td>AWUC</td>
<td>80</td>
<td>.010</td>
<td>.014</td>
<td>.57</td>
<td>-.15</td>
<td>.53</td>
</tr>
</tbody>
</table>

*Note. AWUS = Academic word use for simple task under pre-task-planning and no-pre-task planning conditions, AWUC = Academic word use for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal*

The results show that providing pre-task planning did not influence the learners’ AWU significantly in either the simple, \( z = -.53, p = .60, r = .05 \), or the complex version of the task, \( z = -.86 \ p = .39, r = .09 \).

A graphical illustration of the means and 95% confidence intervals is presented in Figure 5.7. Learners’ AWU improved as a function of increasing task complexity along the level of reasoning and the number of elements. The results were associated with large effect sizes. Providing pre-task planning, however, did not result in the enhancement of AWU, irrespective of the task complexity level.
5.9. Effects of Task Complexity and Planning Conditions on the Number of Words per Minute as a Measure of Fluency (FL, RQ5)

In order to test the effects of the pre-task planning conditions and increasing task complexity by the level of reasoning and the number of elements on English-as-a-foreign-language (EFL) learners’ number of words per minute as a measure of fluency (FL), a mixed ANOVA and follow-up Paired Samples t tests and Independent Samples t tests were conducted. The assumption of normality was met (see Table 5.36) and Levene’s F test for FLS (number of words per minute for the simple task), $F(1, 78) = 2.35, p = .13$, and for FLC (number of words per minute for the complex task), $F(1, 78) = 1.88, p = .17$, was not significant.

Figure 5.7. Bar chart with learners’ AWU means and 95% confidence intervals across four conditions
Table 5.32. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for FL

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>M/SD</th>
<th>n/n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLSNP</td>
<td>40</td>
<td>.20</td>
<td>.84</td>
<td>.16</td>
<td>-.33</td>
<td>9.25/2.59</td>
<td>n</td>
</tr>
<tr>
<td>FLCNP</td>
<td>40</td>
<td>.20</td>
<td>.73</td>
<td>-.17</td>
<td>-.26</td>
<td>7.71/1.98</td>
<td>n</td>
</tr>
<tr>
<td>FLSP</td>
<td>40</td>
<td>.02*</td>
<td>.09</td>
<td>.62</td>
<td>-.14</td>
<td>10.49/1.95</td>
<td>n</td>
</tr>
<tr>
<td>FLCP</td>
<td>40</td>
<td>.07*</td>
<td>.55</td>
<td>.37</td>
<td>-.43</td>
<td>8.63/1.59</td>
<td>n</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, FLSNP = Fluency for simple task under no-pre-task planning condition, FLCNP = Fluency for complex task under no-pre-task planning condition, FLSP = Fluency for simple task under pre-task-planning condition, FLCP = Fluency for complex task under pre-task-planning condition

The descriptive statistics associated with the learners’ FL across four conditions are reported in Table 5.32. As can be seen, simple-pre-task planning, simple-no-pre-task planning, complex-pre-task planning, and complex no-pre-task-planning conditions were associated with the numerically highest to lowest mean level of learner’s FL (M = 10.49, SD = 1.95), (M = 9.25, SD = 2.59), (M = 8.63, SD = 1.59), and (M = 7.71, SD = 1.98), respectively.

The mixed ANOVA yielded significant main effects for increasing task complexity along the resource-directing dimension (increasing the level of reasoning and the number of elements), F(1, 78) = 63.64, p < .001, partial eta-squared ($\eta^2_p$) = .45, and for providing pre-task planning, F(1, 78) = 7.02, p < .001, partial eta-squared ($\eta^2_p$) = .08. However, the interaction effect, F(1, 78) = .58, p = .45, partial eta-squared ($\eta^2_p$) = .007, was not significant. These findings indicate that reducing task complexity along the resource-directing dimension (collapsing across the planning conditions) and providing pre-task-planning (collapsing across the resource-directing dimension) enhanced the learners’ FL significantly. Nevertheless, the learners’ FL on the simple versus the complex versions of the task was not contingent upon planning conditions.
To test the differences between the learners’ production on the simple \((M = 9.25, SD = 2.59)\) compared to the complex version of the task \((M = 7.71, SD = 1.98)\) under the no-pre-task-planning condition, and on the same simple \((M = 10.49, SD = 1.95)\) versus the same complex task \((M = 8.63, SD = 1.59)\) under the pre-task-planning condition, two Paired Samples \(t\) tests were performed. The assumption of normally distributed difference scores was tested and considered to be met (see Table 5.33).

### Table 5.33. Normality Test Results for Paired Samples \(t\) Tests for FL

<table>
<thead>
<tr>
<th>Variables</th>
<th>(N)</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLNP(S-C)</td>
<td>40</td>
<td>.09</td>
<td>.98</td>
<td>.04</td>
<td>.12</td>
<td>.73</td>
</tr>
<tr>
<td>FLP(S-C)</td>
<td>40</td>
<td>.09</td>
<td>.97</td>
<td>-.14</td>
<td>-.33</td>
<td>.73</td>
</tr>
</tbody>
</table>

*Note.* (S-C) = (the score on the simple task – the score on the complex task), FLNP = Number of words per minute under no-pre-task planning condition, FLP = Number of words per minute under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The Paired Samples \(t\) tests and Independent Samples \(t\) tests results are presented in Table 5.35. As can be seen from the table, under both the no-pre-task planning condition, \(t(39) = 4.60, p < .001\), and the pre-task planning condition, \(t(39) = 7.03, p < .001\), as a function of increasing task complexity, the learners’ FL deteriorated significantly. The results for the no-pre-task planning and the pre-task planning conditions were associated with large effect sizes, \(d = .72\) and \(d = 1.11\), respectively.

To examine the differences between the learners’ FL on the simple version of task under the pre-task planning condition \((M = 10.49, SD = 1.95)\) versus the no-pre-task planning condition \((M = 9.25, SD = 2.59)\) and also the difference in their FL on the complex version of the task under the no-pre-task planning condition \((M = 7.71, SD = 1.98)\) and on the same task under the pre-task planning condition \((M = 8.63, SD = 1.59)\) under the no-pre-task-planning condition, two paired samples \(t\) tests were performed. The assumption of normally distributed difference scores was tested and considered to be met (see Table 5.33).
1.59), two Independent Samples $t$ tests were performed. The assumption of normality was examined and considered to be approximately met (see Table 5.34).

### Table 5.34. Normality Test Results for Independent Samples $t$ Tests for FL

<table>
<thead>
<tr>
<th>Variables</th>
<th>$N$</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Stat.</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLS</td>
<td>80</td>
<td>.06</td>
<td>.69</td>
<td>.07</td>
<td>-.04</td>
<td>.53</td>
</tr>
<tr>
<td>FLC</td>
<td>80</td>
<td>.09</td>
<td>.21</td>
<td>-.09</td>
<td>.07</td>
<td>.53</td>
</tr>
</tbody>
</table>

*Note.* FLS = Number of words per minute for simple task under pre-task-planning and no-pre-task planning conditions, FLC = Number of words per minute for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat. = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

As displayed in Table 5.35, providing pre-task planning increased learners’ FL significantly on the simple, $t(78) = 2.42, p < .05, d = .54$, and the complex tasks, $t(78) = 2.29, p < .05, d = .51$. The results were associated with medium effect sizes.

### Table 5.35. Paired and Independent Samples $t$ Tests Results for FL

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$T$</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL-SP</td>
<td>40</td>
<td>10.49</td>
<td>1.95</td>
<td>4.60</td>
<td>39</td>
<td>.000</td>
<td>1.11</td>
</tr>
<tr>
<td>FL-CP</td>
<td>40</td>
<td>8.63</td>
<td>1.59</td>
<td></td>
<td></td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td>FL-SNP</td>
<td>40</td>
<td>9.25</td>
<td>2.58</td>
<td>7.03</td>
<td>39</td>
<td>.000</td>
<td>.54</td>
</tr>
<tr>
<td>FL-CNP</td>
<td>40</td>
<td>7.71</td>
<td>1.98</td>
<td></td>
<td></td>
<td></td>
<td>.51</td>
</tr>
<tr>
<td>FL-SP</td>
<td>40</td>
<td>10.49</td>
<td>1.95</td>
<td>2.42</td>
<td>78</td>
<td>.017</td>
<td>.54</td>
</tr>
<tr>
<td>FL-SNP</td>
<td>40</td>
<td>9.25</td>
<td>2.58</td>
<td></td>
<td></td>
<td></td>
<td>.51</td>
</tr>
<tr>
<td>FL-CP</td>
<td>40</td>
<td>8.63</td>
<td>1.59</td>
<td>2.29</td>
<td>78</td>
<td>.025</td>
<td>.51</td>
</tr>
<tr>
<td>FL-CNP</td>
<td>40</td>
<td>7.71</td>
<td>1.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, $d$ = effect size

A visual representation of the means and 95% confidence intervals is depicted in Figure 5.8. As displayed in the Figure, learners’ FL improved as a function of decreasing task complexity along the reasoning and the number of elements and providing pre-task planning condition, with large and medium effect sizes, respectively.
Figure 5.8. Bar chart with learners’ FL means and 95% confidence intervals across four conditions

5.10. Effects of Task Complexity and Pre-task Planning Conditions on Content of L2 Writing (CN; RQ5)

To study the influence of increasing task complexity by the level of reasoning and the number of elements and the pre-task planning versus the no-pre-task planning conditions on the learners’ CN, the following procedure was used. First, the assumption of the normal distribution for CN scores under each of four conditions (simple task under pre-task planning, complex task under pre-task-planning, simple task under no-pre-task planning, complex task under no-pre-task-planning conditions) was evaluated and found not to be met (see Table 5.36).
As the assumption of normality was violated, instead of mixed ANOVA, other statistical tests were performed to test the impact of task complexity along the resource-directing dimension on CN (within groups analysis) and along the resource-dispersing dimension (between groups analysis). For within group analysis, the assumption of normally distributed difference CN scores for Paired Samples *t* test was examined and considered to be met (see Table 5.37).

### Table 5.37. Normality Test Results for Paired Samples *t* Tests for CN

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CNSNP(S-C)</td>
<td>40</td>
<td>.01*</td>
<td>.01*</td>
<td>.88</td>
<td>.37</td>
<td>.04</td>
<td>.73</td>
<td>21.60/1.61</td>
<td>n.n.</td>
</tr>
<tr>
<td>CNP(S-C)</td>
<td>40</td>
<td>.03*</td>
<td>.06*</td>
<td>-.05</td>
<td>.37</td>
<td>.40</td>
<td>.73</td>
<td>n</td>
<td></td>
</tr>
</tbody>
</table>

Note. (S-C) = (the score on the simple task – the score on the complex task), CNSNP = content score under no-pre-task planning condition, CNP = content under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

Two Paired Samples *t* tests show that the learners’ CN scores increased significantly in the complex (*M* = 22.60, *SD* = 1.44) compared to the simple version of the task (*M* = 21.60, *SD* = 1.61) under the no-pre-task planning condition, *t*(39) = -4.47, *p* < .001, *d* = .70, (see Table 5.38). The same trend was found for the learner’s CN on the simple (*M* = 22.64, *SD* = 1.52) versus the complex version of the task (*M* = 23.35, *SD* = 1.33) under the pre-task planning condition, *t*(39) = -3.43, *p* < .001 *d* = .53 (see Table 5.38). To sum up, as a function of increasing task complexity along the level of
reasoning and the number of elements, learners’ content scores in L2 writing production improved under both conditions. The results were associated with medium and large effect sizes, respectively.

Table 5.38. Paired Samples $t$ Tests Results for CN

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$T$</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN-SP</td>
<td>40</td>
<td>22.64</td>
<td>1.52</td>
<td>-3.43</td>
<td>39</td>
<td>.001</td>
<td>.53</td>
</tr>
<tr>
<td>CN-CP</td>
<td>40</td>
<td>23.35</td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN-SNP</td>
<td>40</td>
<td>21.60</td>
<td>1.61</td>
<td>-4.47</td>
<td>39</td>
<td>.000</td>
<td>.70</td>
</tr>
<tr>
<td>CN-CNP</td>
<td>40</td>
<td>22.60</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, $d$ = effect size

To test the differences between the learners’ CN on the simple version of task under the pre-task planning condition ($M = 22.64, SD = 1.52$) versus the no-pre-task planning condition ($M = 21.60, SD = 1.61$) and also the difference in their CN on the complex version of the task under the no-pre-task planning condition ($M = 22.60, SD = 1.44$) and on the same task under the pre-task planning condition ($M = 23.35, SD = 1.33$), two Independent Samples $t$ tests were run. Prior to running the test, the assumption of normality was evaluated and found to be approximately satisfied (see Table 5.39).

Table 5.39. Normality Test Results for Independent Samples $t$ Tests for CN

<table>
<thead>
<tr>
<th>Variables</th>
<th>$N$</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sig.</td>
<td>Sig.</td>
<td>Stat.</td>
<td>Std. Er.</td>
<td>Stat.</td>
</tr>
<tr>
<td>CNS</td>
<td>80</td>
<td>.01*</td>
<td>.01*</td>
<td>.44</td>
<td>.26</td>
<td>-.71</td>
</tr>
<tr>
<td>CNC</td>
<td>80</td>
<td>.02*</td>
<td>.01*</td>
<td>.19</td>
<td>.26</td>
<td>-.56</td>
</tr>
</tbody>
</table>

Note. CNS = content for simple task under pre-task-planning and no-pre-task planning conditions, CNC = content for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal
As shown in Table 5.40, providing pre-task planning augmented learners’ CN scores significantly on both the simple, \( t(78) = 2.95, p < .05, d = .66 \), and the complex task, \( t(78) = 2.41, p < .05, d = .54 \). The results were associated with medium effect sizes.

**Table 5.40.** Independent Samples \( t \) Tests Results for CN

<table>
<thead>
<tr>
<th></th>
<th>( N )</th>
<th>( M )</th>
<th>( SD )</th>
<th>( T )</th>
<th>( df )</th>
<th>Sig.(2-tailed)</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN-SP</td>
<td>40</td>
<td>22.64</td>
<td>1.52</td>
<td>2.95</td>
<td>78</td>
<td>.004</td>
<td>.66</td>
</tr>
<tr>
<td>CN-SNP</td>
<td>40</td>
<td>21.60</td>
<td>1.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN-CP</td>
<td>40</td>
<td>23.35</td>
<td>1.33</td>
<td>2.41</td>
<td>78</td>
<td>.018</td>
<td>.54</td>
</tr>
<tr>
<td>CN-CNP</td>
<td>40</td>
<td>22.60</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, \( d \) = effect size*

A graphical illustration of the means and 95% confidence intervals is presented in Figure 5.9. As depicted in the Figure, learners’ CN enhanced as a function of both increasing task complexity along the reasoning and the number of elements and providing pre-task planning, with large and medium effect sizes, respectively.

**Figure 5.9.** Bar chart with learners’ CN means and 95% confidence intervals across four conditions
### 5.11. Effects of Task Complexity and Pre-task Planning Conditions on L2 Writing Organisation (WORG; RQ5)

To test the impact of the pre-task planning conditions and increasing task complexity along the level of reasoning and the number of elements on English-as-a-foreign-language (EFL) learners’ WORG, a mixed ANOVA and follow-up Paired Samples $t$ tests and Independent Samples $t$ tests were conducted. First, the assumption of normality for mixed ANOVA were evaluated and considered to be approximately met. Furthermore, the assumption of homogeneity of variances was tested and satisfied by Levene’s $F$ test for WORGS (writing organisation for the simple task), $F(1, 78) = .10, p = .74$, and for WORGC (writing organisation for the complex task), $F(1, 78) = .57, p = .45$.

#### Table 5.41. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for WORG

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WORGSNP</td>
<td>40</td>
<td>.01*</td>
<td>.02*</td>
<td>.12</td>
<td>.37</td>
<td>-.84</td>
<td>.73</td>
<td>14.57/1.35</td>
<td>n</td>
</tr>
<tr>
<td>WORGCNP</td>
<td>40</td>
<td>.01*</td>
<td>.01*</td>
<td>.06</td>
<td>.37</td>
<td>-.93</td>
<td>.73</td>
<td>14.47/1.21</td>
<td>n</td>
</tr>
<tr>
<td>WORGSP</td>
<td>40</td>
<td>.01*</td>
<td>.01*</td>
<td>.23</td>
<td>.37</td>
<td>-.11</td>
<td>.73</td>
<td>14.77/1.35</td>
<td>n</td>
</tr>
<tr>
<td>WORGCP</td>
<td>40</td>
<td>.01*</td>
<td>.01*</td>
<td>.33</td>
<td>.37</td>
<td>-.37</td>
<td>.73</td>
<td>15.37/1.37</td>
<td>n</td>
</tr>
</tbody>
</table>

*Note.* Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, WORGSNP = Writing organisation for simple task under no-pre-task planning condition, WORGCNP = Writing organisation for complex task under no-pre-task planning condition, WORGSP = Writing organisation for simple task under pre-task-planning condition, WORGCP = Writing organisation for complex task under pre-task-planning condition.

The descriptive statistics associated with the learners’ WORG across four conditions are reported in Table 5.41. As can be seen, complex-pre-task planning, simple-pre-task planning, simple-no-pre-task planning, and complex no-pre-task-planning conditions were associated with the numerically highest to lowest mean level of learner’s WORG ($M = 15.37$, $SD = 1.37$), ($M = 14.77$, $SD = 1.35$), ($M = 14.57$, $SD = 1.35$), and ($M = 14.47$, $SD = 1.21$), respectively.
The mixed ANOVA did not yield a significant main effect for increasing task complexity along the resource-directing dimension (increasing the level of reasoning and the number of elements), $F(1, 78) = 2.54, p = .11$, partial eta-squared ($\eta^2_p$) = .03. However, the effect of providing pre-task planning, $F(1, 78) = 4.78, p = .03$, partial eta-squared ($\eta^2_p$) = .05, and the interaction effect, $F(1, 78) = 4.99, p = .02$, partial eta-squared ($\eta^2_p$) = .06, were both significant. These results indicate that increasing task complexity along the resource-directing dimension (collapsing across the planning conditions) did not improve learners’ WORG. Nonetheless, providing pre-task-planning (collapsing across the resource-directing dimension) enhanced the learners’ WORG significantly. Additionally, the learners’ WORG on the simple versus the complex versions of the tasks was contingent upon planning conditions.

To further explore the nature of learners’ WORG across four conditions, the mixed ANOVA analyses were followed up by Paired Samples t tests for within-groups and Independent Samples t tests for between-groups conditions. For the learners’ production on the simple ($M = 14.57, SD = 1.35$) compared to the complex version of the task ($M = 14.47, SD = 1.21$) under the no-pre-task-planning condition, and on the same simple ($M = 14.77, SD = 1.35$) versus the same complex task ($M = 15.37, SD = 1.37$) under the pre-task-planning condition, two Paired Samples t tests were run. The assumption of normality distributed difference scores for Paired Samples t tests was tested and found to be met (see Table 5.42).
Table 5.42. Normality Test Results for Paired Samples \( t \) Tests for WORG

<table>
<thead>
<tr>
<th>Variables</th>
<th>( N )</th>
<th>Ko-Sm Sig.</th>
<th>Sh-W Sig.</th>
<th>Skewness Stat.</th>
<th>Kurtosis Stat.</th>
<th>n./n.n.</th>
</tr>
</thead>
</table>
| WORG
| NBP(S-C)        | 40     | .001*      | .001*     | .06 .37        | -.75 .73       | n       |
| WORGP(S-C)      | 40     | .01*       | .03*      | .33 .37        | -.51 .73       | n       |

Note. (S-C) = (the score on the simple task – the score on the complex task), WORGNP = writing organisation score under no-pre-task planning condition, WORGP = writing organisation score under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

Paired Samples \( t \) tests and Independent Samples \( t \) tests results are reported in Table 5.44. As can be seen from the table, two Paired Samples \( t \) tests show under the no-pre-task planning condition, \( t(39) = .10, p = .67, d = .06 \), learners’ WORG did not vary significantly. However, under pre-task planning condition, \( t(39) = -.2.88, p = .006, d = .45 \), learners’ WORG improved significantly.

The differences between the learners’ WORG on the simple version of task under the pre-task planning condition (\( M = 14.77, SD = 1.35 \)) versus the no-pre-task planning condition (\( M = 14.57, SD = 1.35 \)) and also the difference in their WORG on the complex version of the task under the no-pre-task planning condition (\( M = 14.47, SD = 1.21 \)) and on the same task under the pre-task planning condition (\( M = 15.37, SD = 1.37 \)) were examined via two Independent Samples \( t \) tests. Prior to performing the tests, the assumption of normality was evaluated and considered to be approximately satisfied (see Table 5.43).
As presented in Table 5.44, providing pre-task planning increased learners’ WORG significantly on the complex task, $t(78) = 3.10$, $p = .003$, $d = .70$, but had no significant effect on the simple task, $t(78) = .66$, $p = .51$, $d = .15$. The result for the complex task was associated with a medium effect size.

A visual illustration of the means and 95% confidence intervals is shown in Figure 5.10. As depicted in the Figure, learners’ WORG improved as a function of the interaction effect of increasing task complexity along the reasoning and the number of elements and providing pre-task planning.
5.12. Effects of Task Complexity and Pre-task Planning Conditions on L2 Writing Quality (WQ; RQ5)

A mixed ANOVA and follow-up Paired Samples $t$ tests and Independent Samples $t$ tests were performed to investigate the effect of pre-task planning conditions and increasing task complexity along the level of reasoning and the number of elements on English-as-a-foreign-language (EFL) learners’ WQ. Firstly, the assumptions for the mixed ANOVA were examined and considered to be approximately satisfied (see Table 5.45). Additionally, Levene’s $F$ test revealed that the assumption of homogeneity of variances for WQS (writing quality for the simple task), $F(1, 78) = .92, p = .34$, and for WQC (writing quality for the complex task), $F(1, 78) = .53, p = .46$, was met.
Table 5.45. Normality Test Results for Mixed-ANOVA and Descriptive Statistics for WQ

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>M/SD</th>
<th>n/n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WQSNP</td>
<td>40</td>
<td>.06</td>
<td>.11</td>
<td>.66</td>
<td>.37</td>
<td>73.48</td>
<td>4.24</td>
</tr>
<tr>
<td>WQCNP</td>
<td>40</td>
<td>.09</td>
<td>.36</td>
<td>.49</td>
<td>.37</td>
<td>73.92</td>
<td>4.35</td>
</tr>
<tr>
<td>WQSP</td>
<td>40</td>
<td>.20</td>
<td>.59</td>
<td>.42</td>
<td>.37</td>
<td>76.57</td>
<td>5</td>
</tr>
<tr>
<td>WQCP</td>
<td>40</td>
<td>.20</td>
<td>.40</td>
<td>.36</td>
<td>.37</td>
<td>77</td>
<td>4.83</td>
</tr>
</tbody>
</table>

Note. Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal, WQSNP = Writing quality for simple task under no-pre-task planning condition, WQCNP = Writing quality for complex task under no-pre-task-planning condition, WQSP = Writing quality for simple task under pre-task-planning condition, WQCP = Writing quality for complex task under pre-task-planning condition

The descriptive statistics for the learners’ WQ across four conditions are reported in Table 5.45. As demonstrated in the Table, the learners achieved the numerically highest to lowest mean WQ under complex-pre-task planning ($M = 77$, $SD = 4.83$), simple-pre-task planning ($M = 76.57$, $SD = 5$), complex no-pre-task-planning planning ($M = 73.92$, $SD = 4.35$), and simple-no-pre-task ($M = 73.48$, $SD = 4.24$) conditions, respectively.

The mixed ANOVA yielded significant main effects for increasing task complexity along the resource-directing dimension (increasing the level of reasoning and the number of elements), $F(1, 78) = 11.93$, $p = .001$, partial eta-squared ($\eta^2_p$) = .13, and for providing pre-task planning, $F(1, 78) = 9.07$, $p = .003$, partial eta-squared ($\eta^2_p$) = .10, but the interaction effect, $F(1, 78) = .51$, $p = .45$, partial eta-squared ($\eta^2_p$) = .006, was not significant, indicating that learners benefited from increasing task complexity and pre-task planning in improving their WQ. However, the impact of task complexity on learners’ WQ was not influenced by pre-task planning conditions.
To further examine the effect of task complexity on WQ, two Wilcoxon Signed Rank tests were performed, as the assumption of normally distributed difference scores for Paired Samples $t$ tests was evaluated and considered not to be met (see Table 5.46).

**Table 5.46.** Normality Test Results for Paired Samples $t$ Tests for WQ

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WQNP(S-C)</td>
<td>40</td>
<td>.000</td>
<td>.000</td>
<td>1.22</td>
<td>.53</td>
<td>.73</td>
</tr>
<tr>
<td>WQP(S-C)</td>
<td>40</td>
<td>.000</td>
<td>.001</td>
<td>-1.96</td>
<td>-.91</td>
<td>.73</td>
</tr>
</tbody>
</table>

*Note. (S-C) = (the score on the simple task – the score on the complex task), WQNP = writing quality score under no-pre-task planning condition, WQP = writing quality score under pre-task planning condition, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal*

The learners’ WQ improved significantly in the complex task ($M = 77, SD = 4.83$) compared to the simple version of the task ($M = 76.57, SD = 5$) under the pre-task planning condition, $z = 1.99, p = .047, r = .22$. The same trend was found for the learner’s WQ on the simple ($M = 73.48, SD = 4.24$) versus the complex versions of the task ($M = 73.92, SD = 4.35$) under the no-pre-task planning condition, $z = 2.68, p = .007, r = .30$, (see Table 5.48). These results indicate that increasing task complexity along the level of reasoning and the number of elements was conducive to the learners’ WQ in L2 writing production.

As regards the impact of pre-task planning, as the normality of assumption for Independent Samples $t$ Test was examined and concluded to be violated (see Table 5.47), two Mann-Whitney $U$ tests were performed to test the effects of providing pre-task planning on the learners’ WQ in the simple and complex versions of the task.
Table 5.47. Normality Test Results for Independent Samples $t$ Tests for WQ

<table>
<thead>
<tr>
<th>Variables</th>
<th>$N$</th>
<th>Ko-Sm</th>
<th>Sh-W</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>n./n.n.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sig.</td>
<td>Sig.</td>
<td>Stat.</td>
<td>Std. Er.</td>
<td>Stat.</td>
</tr>
<tr>
<td>WQS</td>
<td>80</td>
<td>.022*</td>
<td>.06</td>
<td>.57</td>
<td>.27</td>
<td>.49</td>
</tr>
<tr>
<td>WQC</td>
<td>80</td>
<td>.20</td>
<td>.23</td>
<td>.44</td>
<td>.27</td>
<td>-.02</td>
</tr>
</tbody>
</table>

Note. WQS = writing quality for simple task under pre-task-planning and no-pre-task planning conditions, WQC = writing quality for complex task under pre-task-planning and no-pre-task planning conditions, Ko-Sm = Kolmogorov-Smirnov, Sh-W = Shapiro-Wilk, Stat = Statistics, Sig. = Significant, Std. = Standard, Er. = Error, n. = Normal, n.n. = Non-normal

The findings are indicative of the facilitative role of providing pre-task planning in learners’ WQ in both the simple, $z = 3.02$, $p = .002$, $r = .38$, and the complex versions of the task, $z = 2.89$, $p = .004$, $r = .32$.

Table 5.48. Wilcoxon Signed Rank Test and Mann-Whitney $U$ Tests Results for WQ

<table>
<thead>
<tr>
<th>Variables</th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$Z$</th>
<th>$df$</th>
<th>Sig.(2-tailed)</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WQ-SP</td>
<td>40</td>
<td>76.57</td>
<td>5</td>
<td>1.99</td>
<td>39</td>
<td>.047</td>
<td>.22</td>
</tr>
<tr>
<td>WQ-CP</td>
<td>40</td>
<td>77</td>
<td>4.83</td>
<td>2.68</td>
<td>39</td>
<td>.007</td>
<td>.30</td>
</tr>
<tr>
<td>WQ-SNP</td>
<td>40</td>
<td>73.48</td>
<td>4.24</td>
<td>3.02</td>
<td>78</td>
<td>.002</td>
<td>.38</td>
</tr>
<tr>
<td>WQ-CNP</td>
<td>40</td>
<td>73.92</td>
<td>4.35</td>
<td>2.89</td>
<td>78</td>
<td>.004</td>
<td>.32</td>
</tr>
<tr>
<td>WQ-SP</td>
<td>40</td>
<td>76.57</td>
<td>5</td>
<td>3.02</td>
<td>78</td>
<td>.002</td>
<td>.38</td>
</tr>
<tr>
<td>WQ-SNP</td>
<td>40</td>
<td>73.47</td>
<td>4.24</td>
<td>2.89</td>
<td>78</td>
<td>.004</td>
<td>.32</td>
</tr>
<tr>
<td>WQ-CNP</td>
<td>40</td>
<td>73.92</td>
<td>4.35</td>
<td>2.89</td>
<td>78</td>
<td>.004</td>
<td>.32</td>
</tr>
</tbody>
</table>

Note. SP = simple task under pre-task-planning condition, CP = complex task under pre-task-planning condition, SNP = simple task under no-pre-task-planning condition, CNP = complex task under no-pre-task-planning condition, sig = significant, $r$ = effect size

A graphical representation of the means and 95% confidence intervals is presented in Figure 5.11. As depicted in the Figure, learners’ WQ improved as a function of increasing task complexity along the reasoning and the number of elements and providing pre-task planning.
Figure 5.11. Bar chart with learners’ WQ means and 95% confidence intervals across four conditions

5.13. Summary of the Findings for Study 2

5.13.1. Effects on Syntactic Complexity

Increasing task complexity and providing pre-task planning enhanced syntactic complexity in L2 students’ writing in terms of SUB. The impact on other measures of writing syntactic complexity was not significant.

5.13.2. Effects on Accuracy

The learners’ accuracy deteriorated significantly with increasing task complexity. To elaborate, the effect on error-free t-units per t-unit was significant with a negligible effect size, while for the number of errors per t-unit the effect size was large. In addition, increasing task complexity under both the pre-task and no-pre-task planning conditions led to a significant decline in learners’ accuracy. Additionally, the effects
of pre-task planning and the interaction of the two independent factors did not cause any significant change in learners’ L2 writing accuracy. For both the simple and complex versions of the task, the effect of providing pre-task planning was not significant.

5.13.3. Effects on Lexical Diversity

Regarding the simultaneous impact of increasing task complexity and pre-task planning conditions (presence vs. absence) on the learners’ LD and AWU in writing task performance, there was no significant effect for LD under any condition. However, the learners’ AWU significantly improved in the complex task performance under both the pre-task and no-pre-task planning conditions. Further, providing pre-task planning did not effect any significant change in learners’ LD and AWU in either the simple or complex task.

5.13.4. Effects on Fluency

Increasing task complexity had an adverse significant effect on fluency in the simple and complex task under both the pre-task and no-pre-task planning conditions. Additionally, mixed ANOVA yielded negative effects for increasing task complexity and absence of pre-task planning conditions, but the interaction effect did not reach statistical significance. To synthesise, providing pre-task planning boosted the learners’ L2 writing fluency. Moreover, the lowest fluency in L2 writing production was found in the complex task performance under the no-pre-task planning condition.
5.13.5. Effects on Content

Writing content scores increased as a function of increasing task complexity; the same results were obtained for increasing task complexity for the learners’ L2 writing scores under both the pre-task and no-pre-task conditions in Study 2. Moreover, the availability of pre-task planning further improved L2 writing content scores. The highest scores were obtained in the complex task performance under the pre-task planning condition.

5.13.6. Effects on Organisation

Significant results were found for providing pre-task planning and for interaction effects. However, the main effect for increasing task complexity was not significant. Further analyses revealed that the effect of task complexity on writing organisation scores reached positive significance under the planning condition only, and the availability of pre-task planning resulted in the improvement in organisation scores only on the complex task performance.

5.13.7. Effects on Text Quality

In Study 2, learners’ text quality was enhanced both under the pre-task and no-pre-task planning conditions as a function of increasing task complexity. Furthermore, desirable effects were obtained for text quality as a consequence of providing pre-task planning on both the simple and complex task performance. The highest text quality was achieved in the complex writing task under the pre-planning task condition.
Chapter 6.

Discussion of Main Findings

6.1. Chapter Overview

The main objective of this PhD thesis was to inquire into the potential effects of task complexity and pre-task planning on, and the mediating role of writing motivation and anxiety in, learners’ writing production in L2 English. In this chapter, the results of Study 1 and then those of Study 2 with regard to each dimension of students’ writing production are discussed in relation to the predictive power of the Cognition Hypothesis (CH) and the Trade-Off Hypothesis (TOH) to explore the extent to which the results confirm the predictions of the two competing hypotheses and also that of Attentional Funnelling.

6.2. Effects on Syntactic Complexity

To capture the effects of task complexity on L2 writing syntactic complexity, a multidimensional conceptualisation of the notion of syntactic complexity was adopted. Task complexity might influence each dimension of writing syntactic complexity differently, as each is believed to be the distinctive feature of L2 writing syntactic complexity at different levels of students’ interlanguage development. Three measures of L2 writing syntactic complexity are used in this study. These are: Mean Length of Clauses (MLC) as a measure of sub-clausal complexity, subordination (SUB; subordinate clauses per clause) as a measure of clausal complexity, and phrasal coordination (PCOR). These three dimensions of complexity
are believed to be the distinctive features of advanced, intermediate, and beginner levels syntactic complexity, respectively (Norris & Ortega, 2009).

The results of Study 1 indicated that increasing task complexity effected no significant changes in MLC and PCOR. However, the impact on SUB was significant. These findings contradict those of previous studies (e.g., Frear & Bitchener, 2015; Kuiken & Vedder, 2007, 2008, 2012) and lend support to Robinson’s (2001a, 2005, 2007a) hypothesis that increasing task complexity along the level of reasoning and the number of elements will increase the use of subordination. These contradictory findings might be due to the task type used. In prior research (e.g., Frear & Bitchener, 2015; Kuiken & Vedder, 2007, 2008, 2012) letter-writing tasks and in my study argumentative writing tasks were used. My findings also support Norris and Ortega’s (2009) conviction that multidimensional, developmentally sensitive measures of L2 writing syntactic complexity should be employed to measure L2 writing syntactic complexity.

The results revealed improvements in upper-intermediate learners’ clausal complexity as measured via subordinate clauses per clause and no significant impact on sub-clausal complexity. Clausal complexity and sub-clausal complexity are shown to be the distinctive features of the upper-intermediate and advanced levels of L2 writers’ text, respectively. These findings might indicate that increasing task complexity may afford opportunities for consolidating learners’ interlanguage system rather than stretching their interlanguage system to the next level of syntactic development in L2 writing. This finding highlights the significance of Skehan’s (2014) differentiation of acquisition in terms of development (e.g., learning new
forms) and control (e.g., accurate and fluent use of previously acquired forms). However, as participating learners’ L2 writing syntactic complexity was not measured prior to task performance, and a general proficiency test and holistic writing scores were used to determine learners’ proficiency level, such interpretation should be treated cautiously. Future research should identify learners’ syntactic development level prior to task performance to provide more insights into whether increasing task complexity expands learners’ interlanguage system. Using other measures of complexity such as syntactic variety and sophistication, and other acquisitional timing measures might show different results as well.

Moreover, it could be argued that these findings might be due to the measures used for capturing advanced forms of L2 writing syntactic complexity. This appears not to be warranted, as MLC is shown to measure advanced writing syntactic complexity (Norris & Ortega, 2009). Task-essential environments (Loschky & Bley-Vroman, 1993) or pragmatic requirements of the writing tasks (Bygate, 1999; Ryshina-Pankova & Byrnes, 2013) and personal choice (Pallotti, 2009) could also provide explanations for these findings. Task-essential environments or task requirements, however, seem not to be relevant in this case, as reasoning can be expressed through using subordination, which increases the number of subordinate clauses per clause, and/or sub-clausal complexity, which is achieved through using grammatical metaphors (e.g., nominalisation), phrasal elaboration, and non-finite clauses, that results in lengthening clauses (Norris & Ortega, 2009). Furthermore, increases in the learners’ use of subordination rather than advanced forms of syntactic complexity is more likely to be the result of the learners’ response to the task requirements using the distinctive features of their current interlanguage system, hence resulting in
consolidating/practicing their interlanguage to a greater extent in complex task performance rather than expanding their interlanguage system with regard to their L2 writing syntactic complexity. This finding suggests the necessity of providing task-based instruction in order to stretch learners’ interlanguage system in terms of L2 writing syntactic development. One such approach is using in-put providing tasks (Ellis, 2013) to draw learners’ attention to sub-clausal complexity in order to expand their linguistic resources; another potential approach is post-task exploitation (Skehan, 2014) in the form of reformulation of learners’ writing production by using sub-clausal complexity and explicitly drawing leaners’ attention to these features.

As regards the mediating role of affective factors, some dimensions of learners’ anxiety and motivation associated moderately negatively and some moderately positively with the general measure of learners’ syntactic complexity in the complex task performance. This confirms Robinson’s (2011) prediction that the role of individual learner factors will be more clearly manifest in the complex task performance. To elaborate, adaptive approaches to learning, namely, mastery goal \((r = .34)\), self-efficacy \((r = .35)\), task value \((r = .30)\), and attribution internal \((r = .28)\) were moderately positively related to mean length of t-unit (MLT). Maladaptive approaches to learning, namely, avoidance goal \((r = -.33)\), performance goal \((r = -.35)\), and external attribution \((r = -.36)\) and one dimension of anxiety, avoidance behaviour \((r = -.28)\), each had a significant moderate negative correlation with MLT.

These findings are consistent with prior studies, which researched the relationship between anxiety and writing performance (Atay & Kurt, 2006; Cheng, Howitz, & Shallert, 1999) and corroborate the necessity of using multidimensional measures of
affective factors and learners’ performance to capture the effects of individual learner factors on learners’ performance (Cheng, 2004; Steinberg & Horwitz, 1986). More importantly, these findings highlight the significance of guiding learners to take adaptive approaches to learning to be able to engage in and sustain their engagement with writing, specifically in performing cognitively complex L2 writing tasks.

For the simultaneous impact of increasing task complexity and pre-task planning conditions (presence vs. absence) on the syntactic complexity of the L2 writing production, there was no significant interaction effect. Increasing task complexity and providing pre-task planning enhanced L2 writing syntactic complexity in terms of SUB. The impact on other measures of writing syntactic complexity was not significant. The findings of Study 2 substantiate Robinson’s (2001a, 2005, 2007a) predication and are in line with those of Study 1 regarding the impact of increasing the reasoning demand and the number of elements on L2 writing syntactic complexity.

Providing pre-task planning also augmented learners’ use of subordinate clauses. This result lends support to the CH (Robinson, 2001a, 2011a, 2011b) and is consistent with the findings of prior studies (e.g., Ellis, 2005; Ellis & Yuan, 2004; Friedlander, 1990; Ishikawa, 2007; Johnson et al. 2012; Kroll, 1990). This finding, however, contradicts those of Ong and Zhang’s (2010) study, in which operationalisation of planning was different from that of this study and other previous studies. Moreover, further increases in the learners’ SUB production in the complex task performance under the pre-task planning condition versus the no-pre-task
planning condition echo Robinson’s (2011a, 2011b) prediction of the positive synergistic effects of pre-task planning and task complexity on L2 production.

The favourable effect of pre-task planning on L2 writing syntactic complexity partly substantiates Robinson’s (2011a) hypothesis that pre-task planning reduces procedural demands on attention and leads to positive outcomes for all dimensions of L2 production. This finding can also be accounted for by employing Kellog et al.’s (2013) proposal of attentional funnelling. Kellog et al. (2013) posit that during the conceptualisation stage of prewriting before writers attempt the first draft (i.e., pre-task planning), they funnel their attention mainly to generating and organisation their ideas. The pre-task planning might leave more time and effort (i.e., cognitive resources) to be funnelled to sentence generation and linguistic encoding during the translating stage of writing, which, in turn, might have yielded in improvements in L2 writing syntactic complexity.

As discussed previously, these findings for the upper-intermediate learners might indicate that increasing task complexity and providing pre-task planning in task performance per se, without the provision of task-based pedagogical interventions (e.g., reformulation of L2 learners’ text using advanced L2 writing syntactic forms), increase using more forms aligned with the learners’ proficiency level (i.e., more subordination). Seemingly, these factors prompt learners to use their automatised forms of syntactic complexity to a greater extent and/or provide opportunities for consolidating previously acquired syntactic forms to a greater extent, rather than expanding learners’ syntactic complexity in L2 writing production by eliciting more instances of advanced level syntactic complexity. Robinson (2011a, 2011b) has
suggested that increasing cognitive task complexity will promote L2 interlanguage development. The results of this thesis cast doubt on the applicability of Robinson’s (2011a, 2011b) prediction to L2 writing syntactic complexity development. However, in this study the effect of task performance on writing syntactic complexity was measured based on one complex writing task performance. Performing a large number of complex writing tasks over a period of time may, in reality, result in the acquisition of developmentally advanced syntactic forms via affording opportunities for attempting advanced syntactic forms for form-meaning mapping in response to complex writing task demands and/or through subsequent task-based pedagogical interventions.

6.3. Effects on Accuracy

Two measures were used to examine the effect of task complexity and pre-task planning on, and mediating role of motivation and anxiety in, the absence and presence of errors in learners’ L2 writing production. The number of errors per t-unit was used to measure the presence and the ratio of error-free t-units to the total number of t-units the absence of errors. For Study 1, the results show that increasing task complexity effected significant negative consequences in accuracy, as measured via the number of errors per t-unit and the ratio of error-free t-units to the total number of t-units. Nonetheless, the effect size for the ratio of error-free t-units to the total number of t-units was small; the number of errors per t-unit better echoed the effect. This reflects the potential dependency of prior findings on the measures employed, which might account for some of the inconsistencies in the literature.
These findings are in line with parts of Ruiz-Funes’ (2015) study and contradict those of others (i.e., Ishikawa, 2007; Kuiken & Vedder 2007, 2008, 2011). However, due to the variability in task types, task manipulation, and measures employed in these studies, cross-comparison may well be inappropriate as noted in the Literature Review chapter of this study. Prior research (i.e., Kuiken & Vedder 2007, 2008, 2011) found positive effect on accuracy and no effects on complexity for increasing task complexity in letter-writing tasks, which might be due to the task type employed. Skehan’s (2009) review of the effect of task types on L2 production showed problem-solving tasks lead to improvements in complexity. The findings of my study corroborate Skehan’s findings. The findings of my study could also be due to the influence of the participants’ L1 on their L2 production.

With regard to the mediating role of anxiety and motivation, the dimensions of writing anxiety did not associate significantly with two measures of accuracy, namely, the ratio of error-free t-units to the total number of t-units and the number of errors per t-unit. Similarly, the association between the dimensions of writing motivation and error-free t-units per t-unit was not significant. Conversely, there was a significant moderate positive correlation between self-efficacy and task value with the number of errors per t-unit in the complex task performance, which contradicts prior research findings on the relationship between affective factors and writing performance in general (Atay & Kurt, 2006; Cheng, Howitz, & Shallert, 1999) and might seem puzzling. As these dimensions of writing motivation had also a moderate positive association with L2 writing syntactic complexity in the complex task performance, it shows that those with high self-efficacy and task value used more complex syntactic structures, which, in turn, led to a decline in their accuracy as a
function of increasing task complexity. Some aspects of affective factors associated with writing production in the complex task confirming Robinson’s (2011a, 2011b) hypothesis that the effect of these factors is more influential in complex task performance rather than in its simple counterparts. This also indicates the importance of using multidimensional measures of both affective factors and learners’ performance to detect the potential association between these variables (Cheng, 2004; Steinberg & Horwitz, 1986).

For the second study, the learners’ accuracy deteriorated significantly with increases in task complexity. To elaborate, the effect of task complexity on the ratio of error-free t-units to the total number of t-units was significant with a negligible effect size, while for the number errors per t-unit the effect size was large. In addition, increasing task complexity under both the pre-task and no-pre-task planning conditions led to a significant decline in learners’ accuracy. These findings are consistent with the results of Study 1, indicating that using multiple measures that tap at different aspects of writing production is necessary. Additionally, the effects of the pre-task planning and the interaction of the two independent factors did not cause any significant changes in learners’ L2 writing accuracy. For both simple and complex tasks, the effect of providing pre-task planning was not significant which confirms some prior findings (e.g., Ellis, 2005; Ellis & Yuan, 2004; Gilabert, 2007), notwithstanding differences in the research design of these studies. The finding also lends support to the Ellis and Yuan’s (2004) findings and argument that pre-task planning may not affect the monitoring phase of the L2 writing process.
Arguably, positive changes in syntactic complexity with no adverse effects on accuracy, as a function of pre-task planning, can be interpreted as a favourable effect for accuracy, as increases in complexity typically destabilize learners’ interlanguage and result in a lower accuracy. This finding can also be accounted for via attentional funnelling (Kellog et al., 2013) in pre-task planning, which might have freed more attentional resources for translating and linguistic encoding of ideas generated and organised during the conceptualisation stage (i.e., pre-task planning) of writing production.

From a pedagogical perspective, increases in syntactic complexity in L2 writing with a decline in accuracy can be viewed with an optimistic lens. Such occurrences can provide opportunities for task-based pedagogical interventions, which, in turn, might enhance learners’ linguistic accuracy along with their syntactic complexity. Task-based learning is more aligned with holistic approaches to learning and is considered to be more conducive to transfer of learning as opposed to traditional piecemeal approaches to instruction (Long, 2015).

In view of the findings with regard to syntactic complexity and accuracy, it seems that these two dimensions draw on a single resource pool. Hence, there is a competition for attention between complexity and accuracy, as increases in syntactic complexity were associated with a decline in accuracy. This disconfirms one aspect of Robinson’s (2001a, 2001b, 2011a, 2011b) CH and supports Skehan’s (1998a, 2009) TOH. This finding also reflects Kellog et al.’s (2013) contention that grammatical encoding might draw on verbal working memory (VWM), indicating increases in syntactic complexity might have overloaded VWM that has led to a
decline in accuracy. Nevertheless, due to the multiplicity of factors involved and inconsistencies in the findings in the literature, partially due to the emerging nature of this strand of research and measures used to detect the effect, such interpretation of the findings should be treated cautiously.

Admittedly, concluding that complexity and accuracy compete for attention in L2 writing production requires substantial empirical evidence. Undeniably, there might be dimensions of accuracy that might have been improved, but the measures employed in this study did not capture such an effect. New approaches to measuring accuracy may find positive effects for accuracy as a function of increasing task complexity (see Foster & Wigglesworth, 2016, for Weighted Clausal Ratio as the most recent proposed measure of accuracy and also Wulff & Gries, 2011, for corpus-driven approaches for assessing accuracy). Moreover, differentiating between lexical and syntactic-morphological accuracy (Kuiken & Vedder, 2008) and/or using specific measures of accuracy (Ishikawa, 2007) might have demonstrated a beneficial effect for accuracy. The results might have also been mediated by other factors not explored in this study (see e.g., two recent studies by Frear & Bitchener, 2015 and Ruiz-Funes, 2015).

Principally, these latter two recent studies (Frear & Bitchener, 2015; Ruiz-Funes, 2015) have shown that the effect of task complexity on writing production is mediated by the degree of difference in the cognitive complexity of tasks used and learners’ level of performance. To illustrate, Frear and Bitchener’ (2015) comparison of learners’ syntactic and lexical complexity on two tasks of varying degrees of complexity revealed a significant favourable effect on lexical complexity, confirming
the CH. However, when the researchers added an obviously simple task to the equation, a significant positive effect for lexical complexity and a significant negative impact for syntactic complexity surfaced, which support Skehan’s (1998a, 2009, 2014) TOH. However, the authors argued that their result was more likely due to the incorrect alignment of the tasks employed in their study with the participating learners’ proficiency level and concluded that too complex or too simple tasks might obscure the results.

Moreover, Ruiz-Funes (2015) found that complex task performance yielded higher syntactic complexity and lower accuracy and fluency irrespective of learners’ proficiency. However, when the learners were assigned into high performance versus low performance groups based on their performance on the writing tasks, complex task performance effected positive changes in high performance advanced level learners’ syntactic complexity, accuracy, and fluency simultaneously. As such, further research inquiries are imperative to identify main and mediating factors and conditions leading to lopsided or maximal noninterfering access to attention, which results in simultaneous improvements in complexity and accuracy or one of the two. Skehan (2009) also has called for research into identifying factors and conditions affecting learners’ CALF to be able to interpret L2 performance.

More importantly, in studies yielding simultaneous increases in syntactic complexity and errors as a function of complex task performance, understanding whether the errors originate from attempting new forms or using forms from an acquired system is significant and has implications for both theory and practice. Theoretically, if syntactic complexity improvements coincide with a decline in accuracy as a result of
learners’ trying new and emerging forms in response to task demands; it should not be considered as evidence for the TOH. This might be indicative of lack of underlying knowledge rather than lack of control due to limited attention.

From a pedagogical perspective, declines in accuracy (namely, making more errors) that arise from learners’ attempting new forms can afford opportunities for task-based pedagogical mediations (e.g., corrective feedback), which, in turn, may result in language learning and development. However, if performing complex tasks elicits more production of acquired forms with a lower accuracy, this can be interpreted as a trade-off effect that suggests lack of availability of sufficient attention to attend to form. In this study, it is more likely that increases in L2 writing syntactic complexity in terms of the more use of subordinate clauses per clause, as a function of increasing task complexity, were derived from the learners’ acquired system, and hence higher syntactic complexity and simultaneous lower accuracy can be cautiously interpreted as a trade-off effect.

6.4. Effects on Lexical Complexity

Two measures of lexical complexity were employed to probe the impact of task complexity and pre-task planning on, and the role of motivation and anxiety in, the learners’ L2 writinglexical complexity. D value, as a measure of lexical diversity, and AWU, as measures of advanced level vocabulary use, were employed. For Study 1, the results indicated that increasing task complexity had no significant effect on D value. AWU, however, increased significantly in the complex task performance. The same pattern emerged for the learners’ lexical complexity in the simple versus complex tasks performance under both the pre-task and no-pre task planning.
conditions which confirms prior research findings (Kormos, 2011; Kuiken & Vedder, 2007, 2008, 2011). Additionally, simultaneous increases in syntactic and lexical complexity lend support to the CH, as learners used more advanced academic words in the complex task performance and also encoded their ideas using more complex syntactic forms. Furthermore, despite the insignificant correlation between the SUB and AWU in Study 1, frequency analysis showed simultaneous improvements, not the trade-off effect, were the norm at the individual level. These findings challenge Skehan’s (2009) claim that despite the experimental effects on the dimensions of L2 production, the trade-off effect might be the norm at the individual level and follow-up correlation analysis is needed to demonstrate simultaneous improvements.

As for the mediating role of affective factors, only avoidance behaviour (AB), one dimension of the learner’ writing anxiety, had a significant moderate negative association \( r = -.27 \) with AWU in the complex task performance. The correlation between other dimensions of writing anxiety and lexical complexity was not significant in either the simple or the complex task performance. Additionally, the mediating role of motivation in AWU in writing performance did not reach statistical significance in the simple or complex task. As discussed previously, the findings confirm the negative contribution of anxiety to writing performance (e.g., Atay & Kurt, 2006; Cheng, Howitz, & Shallert, 1999) and lend support to Robinson’s (2011a, 2011b) hypothesis that the mediating role of affective factors might be more clearly exhibited in performing cognitively complex interactive tasks. Furthermore, no mediating role of cognitive anxiety (CA) and somatic anxiety (SA) in the simple or the complex writing task performance, and the significant negative mediating role of avoidance behaviour in the complex writing performance, highlight the role of
engagement in L2 writing production, suggesting the significance of adapting pedagogical mediations that promote learners’ engagement in L2 writing (See Philp & Duchesne, 2016, for discussion of engagement in task-based L2 learning).

Regarding the simultaneous impact of increasing task complexity and pre-task planning conditions (presence vs. absence) on the learners’ LD and AWU in L2 writing task performance, there was no significant effect for LD under any condition. However, the learners’ AWU significantly improved in the complex task performance under both the pre-task and no-pre-task planning conditions, confirming the results of Study 1 and the CH with regard to the effects of the resource-directing dimension of task complexity on L2 complexity. Furthermore, in line with Johnson, Mercado, and Acevedo’s (2012) findings, providing pre-task planning did not effect any significant changes in learners’ LD and AWU in either the simple or complex tasks. This seems to support Pallotti’s (2009) argument that it might be the task requirements and its goal rather than the availability of attentional resources that result in improvements in the complexity of learners’ performance. Nonetheless, as both task complexity and pre-task planning led to increases in syntactic complexity in this study, in some instances availability of pre-task planning, which reduces the procedural demands of the tasks (Robinson, 2011a, 2011b), leads to increases in using more subordinate clauses (i.e., syntactic complexity) that may be the result of attentional funnelling (Kellog et al., 2013).

6.5. Effects on Fluency

The number of words per minute of learners’ writing production was employed to investigate the impact of task complexity and pre-task planning on, and the role of
motivation and anxiety in the learners’ L2 writing fluency. In Study 1, increasing task complexity had a significant adverse effect on fluency. The same trend was found for the learners’ fluency in the simple versus complex tasks under both the pre-task and no-pre task planning conditions in Study 2 that confirms both Robinson’s (2011a, 2011b) and Skehan’s (2009) predictions. Additionally, a mixed ANOVA yielded a main adverse effect for increasing task complexity and absence of pre-task planning conditions, but the interaction effect did not reach statistical significance. To elaborate, providing pre-task planning boosted the learners’ L2 writing fluency in line with some previous research along this line of inquiry (Ellis & Yuan, 2004; Gilabert, 2007; Johnson, Mercado, & Acevedo, 2012), notwithstanding the differences in the task types employed. Moreover, the lowest fluency in L2 writing production was found in the complex task performance under the no-pre-task planning condition that reflects Robinson’s (2011a) prediction of the adverse synergistic effects of increasing task complexity and the unavailability of planning time on fluency in L2 production.

Regarding the mediating role of motivation and anxiety in L2 writing production in simple and complex task performance, no significant association was found between the dimensions of motivation and anxiety and L2 writing fluency. The role of affective factors may be more influential in complex task performance under dialogic task conditions such as interactive tasks as opposed to monologic L2 writing task performance, as Robinson (2011a) posits.
6.6. Effects on Content

In Study 1, writing content scores increased as a function of increasing task complexity; the same results were obtained for increasing task complexity for the learners’ L2 writing scores under both pre-task and no-pre-task conditions in Study 2. Moreover, availability of pre-task planning further improved L2 writing content scores. The highest score was obtained for the complex task performance under the pre-task planning condition, confirming a synergistic effect as predicted by the CH. The CH does not make explicit predictions regarding the effect of task complexity and pre-task planning on the dimensions of L2 writing production other than CALF. However, simultaneous improvements in writing content scores and other dimensions of L2 writing production (e.g., syntactic and lexical complexity) confirm the CH prediction that increasing conceptual demands and reducing procedural demands of the task can result in drawing learners’ attention to the dimensions of L2 production and free-up attentional resources, respectively, and hence improve L2 production.

The favourable results of this study corroborate those of Hamp-Lyons and Mathias (1994). Simultaneous improvements in linguistic encoding and higher-order dimensions of L2 writing production also address Kuiken and Vedder’s (2008) concern that improvements in linguistic dimensions of the L2 production may take attention away from the other higher-order dimensions of L2 production and also Adam et al.’s (2014) uncertainty as to whether improvements in CALF affect task outcomes adversely.

As regards the mediating role of motivation and anxiety in L2 writing content in the simple and complex task performance, no significant correlation was found between
the dimensions of anxiety and L2 writing content score. Nonetheless, there was a significant moderate positive association between mastery goal and L2 writing content scores in the complex task performance, indicating the significance of encouraging learners to embrace adaptive approaches to learning. Other dimensions of motivation and L2 content did not reach statistical significance. Once more the results verify Robinson’s (2011a, 2011b) prediction regarding the influential role of affective factors in complex task performance and also the necessity of adopting multiple measures of affective factors and task performance in studying the role of affective factors in L2 performance.

6.7. Effects on Organisation

In Study 1, increasing task complexity yielded a significant favourable effect for L2 writers’ organisation scores. In line with Révész’s (2011) findings in oral task performance, the correlation between the dimensions of motivation and L2 writing organisation was not significant in either the simple or complex L2 writing task performance. The same results were obtained for the association between dimensions of writing anxiety and writing organisation.

In Study 2, significant main effect was found for providing pre-task planning and there was an interaction effect. However, the main effect for increasing task complexity was not significant. Further analyses revealed that the effect of task complexity reached positive significance under the planning condition only. Pre-task planning resulted in improvement in writing organisation only on the complex task performance. The results of Study 1 and 2 confirm Robinson’s (2011a, 2011b) CH with regard to its predication of improvements in L2 production as a function of
increasing task complexity and synergistic effect of pre-task planning and increasing task complexity. The results are in line with Hamp-Lyons and Mathias (1994) study and also address Kuiken and Vedder’s (2008) and Adam et al.’s (2014) concerns. Improvements in linguistic encoding of the L2 production may not take attention away from higher-order dimensions of L2 production, as simultaneous desirable effects were found for both dimensions of L2 writing production in this study.

The findings of this study can be explained with recourse to Kellog et al. (2013) contention that different dimensions of writing production might draw on different components of working memory. It may even be the case that long-term working memory supplements limited attentional resource of working memory in response to L2 writing task demands. For Study 1, the simultaneous improvements in syntactic and lexical complexity, content, and organisation confirm the CH and Kellog et al.’s findings that different dimensions of writing production may rely on different components of working memory. For Study 2, further improvements in syntactic and lexical complexity, content, organisation, and fluency of writing production can be attributed to the funnelling of attention during phases of L2 writing production, which can result in reduction of procedural demands. However, as increases in syntactic and lexical complexity were associated with a decline in accuracy in Study 1, this result supports the TOH that predicts competition for attention between syntactic complexity and accuracy.

### 6.8. Effects on Text Quality

In Study 1, L2 writers’ writing text quality improved as a result of increasing task complexity in the level of reasoning and the number of elements. The same
favourable trend was found in Study 2, as learners’ text quality improved both under the pre-task and no-pre-task planning conditions as a function of increasing task complexity. Further desirable effects were obtained for text quality as a consequence of providing pre-task planning in both the simple and complex task. The highest text quality was achieved in the complex writing task under the pre-planning task condition. The effect of task complexity was not contingent upon the availability of pre-task planning time, as no significant interaction effect was found in Study 2.

In view of the effects on other dimensions and on writing quality, the findings lend support to Robinson’s (2011a, 2011b) prediction with regard to the possibility of simultaneous enhancements in the dimensions of L2 production as a result of increasing task complexity and synergistic effects of pre-task planning and increasing task complexity. The results also corroborate Hamp-Lyons and Mathias (1994) study in which learners obtained higher scores on the task judged by the expert raters as more difficult. Further, some scholars have convincingly raised issues with measuring only linguistic production in L2 production by arguing that favourable linguistic outcomes might be achieved at the cost of adverse effects for higher-order dimensions of L2 production. Specifically, Pallotti (2009) has posited that it is possible to produce syntactically and lexically complex and accurate structures which might not meet the communicative demands of the task and be pragmatically inappropriate. Adam et al. (2014) and Kuiken and Vedder (2008) have also expressed similar concerns. Simultaneous desirable effects for linguistic encoding in L2 writing production and writing text quality address such concerns and provide support to the possibility of enhancement in L2 writing linguistic production (i.e., CALF) without negative effects for task outcome and functional adequacy.
The mediating role of dimensions of motivation and anxiety did not reach statistical significance in either the simple or complex L2 writing task performance, confirming Révész’s (2011) findings in oral task performance. These results show that the mediating role of ability and affective factors might not be detected by holistic measures of L2 writing production.

As I stated in the previous section, the findings of this study can be interpreted by drawing on Kellog et al.’s (2013) proposal that different dimensions of writing production might employ different components of working memory. It may be the case that limited attentional resources of working memory are supplemented by long-term working memory in response to L2 writing task demands. The simultaneous enhancements in syntactic and lexical complexity, content, organisation, and writing quality in Study 1 confirm the CH and Kellog et al.’s findings that different dimensions of writing production may draw on different components of working memory. Attentional funnelling that might have occurred due to the recursive nature of planning, translating, reading, and editing processes of L2 writing production might have also contributed to simultaneous improvements without overloading WM. Despite simultaneous improvements in other dimensions, negative consequences for accuracy with increases in syntactic complexity on the complex task performance indicate that linguistic encoding might draw on verbal working memory, and increasing task complexity might result in overloading verbal working memory, leading to the trade-off effect as predicted by Skehan’s (1998, 2009) TOH.
Further improvements in syntactic and lexical complexity, content, organisation writing quality, and fluency and no adverse effect for accuracy in L2 writing production in Study 2 can be accounted for by recourse to attentional funnelling. Attentional funnelling might have occurred during phases of L2 writing production afforded by pre-task planning and the recursive nature of L2 writing production, resulting in reduction of procedural demands. All in all, the findings lend partial support to the CH and to the TOH and confirm the potential role of complex argumentative writing task performance in improving L2 writing production and in writing to learn as proposed by Manchon (2014).

It should be noted that the studies cited here supporting or disconfirming the findings of this current study have used different task types, measures of linguistic production, task manipulation, and language proficiency levels that do not correspond to the variables investigated in this study. This is due to the emerging nature of this strand of research within the field of L2 writing. The inconsistency in the findings can also be accounted for by this variability in the operationalisation of the independent, dependent, and the mediating variables investigated (Ruiz-Funes, 2014, 2015). Future research is warranted to inquire into the dimensions of these variables to crystalize prior research findings and provide insight into factors leading to competition for, or multiple access to, attention. Nonetheless, a more important question to be addressed is whether increasing task complexity per se, both in case of attention to one aspect of linguistic production or several aspects, results in eliciting more automatised or partially acquired linguistic forms that are the distinctive features of the learners’ current interlanguage system, or whether it also prompts learners to attempt to use forms that are features of the level beyond the learners’ current interlanguage system.
An alternative interpretation of the findings is also highly possible. Skehan (2009, 2014) interpreted the findings of research into the effect of task characteristics and implementation on L2 oral production by drawing on Levelt’s (1989, 1999) model of L1 speaking and the role of working memory. The same can be applied to the findings of this study by drawing on the Kellog’s (1996) and Kellog et al.’s (2013) model of working memory in writing. It might indeed be the case that increasing task complexity along the level of reasoning demands and the number of elements results in generating complex ideas during the planning stage of writing production. This may lead to using more complex syntactic forms and advanced vocabulary that may consume limited cognitive resources and result in negative consequences for accuracy.

The same may have happened with providing 10-minute pre-task planning except that because of the attentional fuelling effect, due to availability of pre-task planning, attention might have been freed for linguistic encoding during the translating stage (Kellog et al., 2013). As a result, increases in task complexity under the pre-task planning condition, when compared to increasing task complexity under no-pre-task planning condition, result in favourable effects for several dimensions of L2 writing production without leading to negative consequences for accuracy. Clearly, further research is required to test the validity of interpreting the effects of task complexity on L2 writing production with recourse to a theoretical model of working memory.

6.9. Effects of Task Complexity on L2 Writing Task Difficulty

The effect of task complexity on L2 writing task difficulty was measured via 10 teachers and all participating students’ judgements of the complexity of the two tasks
and the students’ rating of the difficulty level of two tasks on a 100-point Likert-type scale. The vast majority of the participating students 92% \( (n = 129) \) and all the teachers found the six-project task more difficult than the three-project task. Further, the students’ rated the complex task as significantly more difficult than the simple one. The following reasons were given by the participants for the higher complexity of the complex versus the simple task: the larger amount of total fund to allocate, the greater number of elements (projects), the requirement for more reasoning to prioritise this number of projects, the competition (trade-off) among the larger number of elements (projects), and the need for more concentration and time and a wider range of vocabulary items.

These findings confirm that the construct of task complexity was operationalised appropriately in this study. This lends support to prior studies that increasing the number of elements and the degree of reasoning increases the complexity of the tasks (Kim, Payant, & Pearson, 2015; Prabhu, 1987; Révész, 2011, 2014; Révész et al., 2015; Robinson, 2001a, 2005). The results also extend the previous findings by revealing that when the elements are closely related and focusing on one element affects the other elements (termed elements trade-off in this study), the task becomes even more cognitively complex than when there is no such effect. This partially results from more internal cognitive comparisons as in Kim et al.’s (2015) study. To elaborate, the three project task in this study is more complex than a task that requires giving directions for moving from point A to B and then to C, also a three-element task, as there is no elements trade-off effect between A, B, and C.
6.10. Summary of the Discussion

The findings of the two studies (i.e., Study 1 and 2) support the CH (Robinson, 2001a, 2001b, 2007b, 2011a, 2011b) as operationalised in the TCF in a number of ways: (a) Increasing the level of reasoning and the number of elements effects in the complexity of the task; (b) increasing task complexity along the level of reasoning and the number of elements leads to the enhancements in syntactic and lexical complexity, content, organisation, and writing quality and has a negative impact on fluency; (c) the mediating role of affective factors is likely to be more evident in complex task performance, and (d) the availability of pre-task planning lessens procedural demands and effects desirable changes in L2 writing production.

However, as the increases in syntactic and lexical complexity in Study 1 were associated with adverse effects for accuracy, the findings afford support to the TOH as well (Skehan, 1998a, 2009). Moreover, improvements in L2 writing syntactic complexity, content, organisation, and writing quality, without adverse effects on accuracy, as a result of 10-minute pre-task planning, corroborate attentional funnelling theory (Kellog et al., 2013) and are in line with the predictions of both the CH and TOH. The findings also support using learners’ self-rating and learners and their teachers’ judgments as reliable procedures for validating task complexity criteria (see also, Révész, 2014; Révész et al., 2015). Furthermore, the results highlight the necessity of using multidimensional measures of motivation and anxiety in order to find the potential mediating role of these factors in L2 writing task performance (Cheng, 2004). Similarly, multidimensional conceptualisation and operationalisation of measures of syntactic complexity (Norris & Ortega, 2009) are supported by the findings.
Chapter 7.

Conclusions, Implications, Limitations, and Recommendations for Further Research

7.1. Overview

In this chapter, the rationales for the study are revisited. Then, the findings of the two studies are summarized. This is followed by the implications of this thesis to theory, methodology, and pedagogy. The chapter concludes with the limitations of the study and recommendations for future research.

7.2. Summary of the Rationales for the Study

Tasks have been assigned a central role in L2 learning and L2 writing pedagogy (Hyland, 2003) due to a number of reasons, including: The potential transferability of the task-based language learning to real-life task performance (Ellis, 2013; Lightbown, 2008; Long 2015), natural acquisitional processes resulting from task performance, and the compatibility of the TBLT with the learners’ developmental level, what Corder (1967) termed ‘built-in syllabus’. The latter provides opportunities for learners to learn the forms that they are developmentally ready to acquire (Ellis, 2013). Understandably, owing to the centrality of tasks in TBLT, research has investigated the impact of task performance on language learning and/or learning opportunities arising from task performance to provide insights into TBLT. In line with this strand of research in L2 writing (e.g., Kuiken & Vedder, 2007, 2008, 2012; Frear & Bitchener, 2015), the effects of writing task design and implementation
features on, and the mediating role of individual factors in, writing production were explored in this study. The main objective of the study was to extend the emerging body of research that aims at identifying factors and conditions that are conducive to writing to learn L2 and learning to write in L2 (Manchón, 2011) in order to provide potential insights into task-based L2 writing pedagogy and assessment.

In order to study the impact of the task complexity, criteria for task complexity should be employed to design tasks of varying degrees of cognitive complexity, which can lead to establishing the validity of such criteria. The validated criteria can be used in designing and developing pedagogical tasks and in sequencing the tasks from simple to complex in order to form a comprehensive task-based writing programme. The criteria can also be employed in designing and developing assessment tasks that are aligned with the learners’ proficiency level and also in designing and developing parallel assessment tasks.

Furthermore, by probing the effects of task design and implementation features on learners’ writing production, insights can be gleaned with regard to factors and conditions that lead to simultaneous or lopsided improvements in the dimensions of L2 production which, in turn, can shed lights on whether learners’ attentional resources are limited or multiple. Such findings can have significant theoretical implications, as the findings can be used to interpret L2 performance and learning. Moreover, by studying the impact of writing task complexity on the linguistic encoding of learners, the potential of the writing tasks performance per se in interlanguage consolidation and/or development and the opportunities that arise for task-based pedagogical mediations can be recognised.
Motivated by the above rationales for investigating the role of task design and implementation features, and individual learner differences in L2 writing production, I designed and conducted two studies. Study 1 addressed the effect of increasing the level of reasoning and the number of elements on, and the mediating role of motivation and anxiety in, L2 writing syntactic complexity, accuracy, lexical complexity, and fluency (CALF), and content, organisation, and writing quality. A simple task (a three-project task) and a complex task (a six-project task) were designed based on the TCF (Robinson, 2011a, 2011b). The validity of the operationalised tasks was confirmed via the learners’ self-rating and 10 teachers and all participating learners’ judgments of the difficulty of the two tasks. In Study 2, the role of task complexity and pre-task planning was investigated. Two groups of participants performed the simple and complex tasks under the pre-task and no-pre-task planning conditions. The same tasks and procedures, which were used in Study 1, were utilised in Study 2 except that the mediating role of motivation and anxiety was not explored in the second study.

7.3. Summary of Main Findings

The results of Study 1 revealed that increasing task complexity led to the enhancement of one dimension of syntactic and one dimension of lexical complexity, had adverse effects on accuracy and fluency, and effected favourable changes in L2 content, organisation, and writing quality. As regards the mediating role of motivation and anxiety, moderate negative associations were found between some dimensions of maladaptive approaches to learning and some dimensions of L2 writing production in the complex writing task performance. In contrast, there were moderate positive correlations between some aspects of adaptive approaches to
learning and some dimensions of L2 writing production in the complex L2 task performance. In Study 2, a similar trend was found for the effect of increasing task complexity along the level of reasoning and the number of elements as in Study 1. With regard to the impact of a 10-minute pre-task planning time on L2 writing production, favourable results were obtained for one dimension of syntactic complexity and fluency, no effects for accuracy and lexical complexity, and significant positive changes for content, organisation, and writing quality.

7.4. Theoretical Implications
Theoretically, these findings mainly support the CH (Robinson, 2001a, 2001b, 2007b, 2011a, 2011b), operationalised in the TCF, in an number of ways: (a) Increasing the level of reasoning and the number of elements results in increasing the cognitive complexity of the task; (b) increasing cognitive task complexity along the level of reasoning and the number of elements results in improvements in syntactic and lexical complexity, content, organization, and writing quality and has adverse effects on fluency; (c) the mediating role of affective factors are more clearly manifest in complex task performance and might be more influential in performing tasks under different task conditions than in tasks with different degrees of cognitive complexity, and (d) providing pre-task planning reduces procedural demands and effects positive changes in L2 performance. However, as the improvements in syntactic and lexical complexity in Study 1 were associated with adverse effects on accuracy, the results lend support to the Trade-Off Hypothesis (TOH) as well (Skehan, 1998a, 2009, 2014). Additionally, improvements in L2 writing production, as a function of pre-task planning, support attentional funnelling theory (Kellog et al., 2013) and are in line with the predictions of both the CH and TOH.
Another theoretically significant issue raised by this study is distinguishing the source of syntactic enhancement and corresponding increases in the number of errors. To illustrate, if learners make more errors as a result of attempting new syntactic structures to meet the demands of the task, this evidence should not be interpreted as a trade-off effect. The errors may originate from the learners’ lack of underlying knowledge of the forms rather than their lack of control over correctly using the forms as a function of unavailability of attention. Additionally, if complex task performance and availability of planning lead to using more instances of already acquired or partially acquired forms that are the distinctive features of the learners’ current proficiency level; it should be interpreted as opportunities for language consolidation or practice rather than interlanguage development in terms of stretching L2 syntactic complexity. Although creating opportunities for consolidating already acquired forms (e.g., L2 subordination) and provision of pedagogical mediations to improve learners’ accuracy are certainly conducive to the enhancement of learners’ interlanguage in terms of control (accuracy and fluency) and may result in learning new linguistic items (e.g., vocabulary), interpreting these developments without differentiating them from expanding L2 syntactic complexity might lead to channelling instruction to accuracy and depriving the learners of the pedagogical mediations that might be necessary for L2 writing syntactic development.

To elaborate, if performing a complex L2 writing task generates more syntactic subordination and simultaneously more errors in students’ writing at the upper-intermediate level, as is the case in this study, and using more subordination is interpreted as syntactic development and corrective feedback is provided on learners’ errors, this might result in the enhancement of the learners’ accuracy and learning
new linguistic items (e.g., new vocabulary and past tense). However, if using more subordination is interpreted as the most advanced syntactic development in L2 writing production, pedagogical mediations aimed at helping the learners to learn developmentally more advanced syntactic forms such as grammatical metaphor (e.g., nominalisation) and non-finite clauses might not be provided. Conversely, if such differentiation is realised, pedagogical mediations can target both syntactic development and accuracy that may lead to the balanced development of linguistic encoding in L2 writing.

Ellis (2013) and Skehan (2014) have proposed approaches to expand learners’ interlanguage. Ellis observes that the type of language that learners use when performing tasks arises in part from the task as workplan (e.g., task type and task design features), in part from task implementation, and also in part from the learners’ developmental level. If task design and implementation features do not lead to the use of the developmentally more advanced forms of language, one pedagogical decision would be using input-providing tasks that require learners to process syntactically advanced forms of language to expand their interlanguage. Ellis (2013) argues that designing such input-proving tasks that necessitate using specific target structures is easier that designing their output-providing counterparts. The complex forms processed during input-providing tasks can provide the learners with the linguistic resources that can be used in subsequent output providing tasks (e.g., writing tasks). These linguistic forms acquired through input-providing tasks can be consolidated via output-prompting tasks if output providing tasks do not lead to stretching learners’ interlanguage. Likewise, in line with Skehan’s proposal of post-task exploitation, other contextualised pedagogical options, such as reformulating
learners’ written output by using advanced L2 writing syntactic forms, can help the learners expand their L2 syntactic complexity to the next level if it is shown that output prompting tasks alone without pedagogical intervention do not expand learners’ interlanguage and only provide opportunities for consolidating learners’ interlanguage.

Additionally, to establish whether a trade-off effect exists, one should first establish that dimensions of L2 production indeed require sufficient attention. To illustrate, if the dimensions of interest rely on using proceduralised knowledge, which requires little attentional resource (Ellis, 2003), simultaneous increases in the dimensions of production should not be interpreted as evidence to disconfirm the TOH. Using newly acquired, but non-proceduralised syntactic forms that concur with a simultaneous decline in accuracy can be indicative of the trade-off effect. Also, errors that arise from using new forms, for which the learners have not acquired the underlying knowledge, should not be considered as evidence for the TOH. Simultaneous enhancements of syntactic and lexical complexity, and even accuracy that occur for the proceduralised forms should not be interpreted as access to multiple resources either because limited attention resources might have been managed successfully.

Furthermore, increases in accuracy should not be considered favourable without considering the syntactic complexity, as learners might resort to using forms already proceduralised via a conscious avoidance strategy. In short, understanding the sources of improvements and declines in L2 production and also the degree of automaticity in L2 use are desirable in interpreting research findings. Better research
evidence for identifying any potential trade-off effect must be obtained through requiring the use of non-proceduralised forms in tasks in which the use of these forms is essential for the successful task completion. In such tasks, if the use of the newly-acquired syntactically complex forms results in more errors, then the trade-off has occurred between accuracy and complexity. If both complexity and accuracy increase, this supports the CH. In cases where accuracy drops as a result of using more instances of forms that are the distinctive features of learners’ current interlanguage system, this drop can be used as evidence for the trade-off effect.

7.5. Methodological Implications

Methodologically, the findings confirm using learners’ self-rating and learners and their teachers’ judgments via open-ended questions as reliable methods for validating the criteria for cognitive task complexity (see also, Révész, 2014; Révész et al., 2015). Moreover, the results highlight the necessity of employing multidimensional measures of motivation and anxiety in order to identify the potential mediating role of these factors in L2 writing task performance (Cheng, 2004). Likewise, multidimensional conceptualisation and operationalisation of measures of linguistic production (Norris & Ortega, 2009) are foreshadowed by the findings, as the effect of cognitive task complexity on L2 writing complexity was captured by one dimension of L2 syntactic and one dimension of L2 lexical complexity. More importantly, developmentally sensitive measures of L2 production should be used to address potential misalignment problems between learners’ developmental level and measures of L2 production and also to identify whether increasing cognitive task complexity stretches learners’ interlanguage system.
7.6. Implications for Writing Pedagogy and Assessment

Five pedagogical implications can be offered based on the findings of this study: (a) Teachers should consider aligning the degrees of the task complexity with their L2 learners’ proficiency by adjusting the level of reasoning and the number of elements required for successful completion of the assigned tasks; (b) teachers can increase the complexity of the tasks along the level of reasoning and the number of elements in order to provide opportunities for using more subordination, to improve lexical complexity, content, organisation, and writing quality, and also to create opportunities for task-based contextualised pedagogical mediations for the development of both complexity and accuracy; (c) teachers can afford further opportunities for learners by allocating sufficient pre-task planning time to cognitively complex task performance in order to increase the use of subordinate clauses and improve fluency and writing quality without resulting in adverse effects for accuracy; (d) teachers should promote adaptive approaches to learning via various pedagogical and assessment approaches, so that learners can initiate and sustain engagement with cognitively demanding tasks of L2 writing, and (e) teachers should employ input providing tasks and post-task exploitation strategies to extend learners’ linguistic resources.

Some suggestions can also be gleaned from the findings of this study for task-based writing assessment. As the complex task performance under the pre-task planning condition elicited the learners’ best writing production without adverse consequences for L2 writing accuracy, in task-based writing assessment, it should be ascertained that the learners are assigned tasks of sufficient complexity with sufficient pre-task planning time. Additionally, as adjusting the level of reasoning and the number of elements
elements in task requirements can lead to varying degrees of cognitive task complexity, this factor should be considered in aligning the writing tasks with learners’ writing developmental level and also in designing and developing parallel writing tasks.

7.7. Limitations

This study relied on quantitative data mainly; hence using a mixed methods design in future studies may provide further insights into the effects of task design and implementation features and pre-task planning on the product of L2 writing and also on processes. No baseline data were collected from English L1 writers; therefore, future research can extend this study by comparing L1 writers’ productions with those of L2 learners to identify the extent to which the observed effects of the task on L2 writing production can be accounted for by task design effect and by L2 learners’ processing capacity. Developmentally sensitive and general measures of L2 writing production were used in this study. Although both the CH and TOH claim that general measures of L2 production will be sensitive to capture the task design and implementation effects, the CH further proposes using specific measures guided by the demands of the tasks employed (Robinson, 2015). Using both general and specific measures of L2 production is desirable in further research.

7.8. Recommendations for Further Research

This study focused on three aspects of Robinson’s CH, namely, the degree of reasoning required, the number of elements involved, and the availability of pre-task planning time. Understandably, one cannot study all aspects of task complexity in a single project. As such, future research will need to investigate the predictions of the
CH with regard to the role of other aspects of the task complexity in second language writing. Specifically, future research should investigate the interaction effect of task complexity along the resource-directing dimension in relation to Here-and-Now versus There-and-Then conditions with the pre-task planning and no-pre-task planning conditions on second language writing production and text quality. In the same study, the modulating role of individual learner difference (ID) factors, including motivational beliefs, anxiety, working memory, and language learning aptitude should be included to further illuminate the interaction impact of task design features and ID factors on SL/FL writing production and text quality. Moreover, the effect of interaction between the resource-directing dimension (including +/-Few elements, +/-Here-and-now, and +/-No reasoning demands) and +/-Prior Knowledge (including familiarity of topic and its predictability, familiarity of discourse features, and familiarity of task), and the modulating effect of ID factors on SL/FL writing production and text quality are also significant aspects to be examined in future research projects.

This study probed into the effect of task complexity along the resource-directing dimension and pre-task planning condition on, and mediating role of affective factors in, L2 writing production based on learners’ individual task performance in the simple and complex versions of the same task. An interesting strand of research will be inquiring into the effects of task complexity and task conditions on the collaborative writing task performance in the simple and complex writing tasks to explore the effects on writing production and processes, and also language and writing learning opportunities that arise from collaborative writing task performance in the simple versus complex writing tasks. Additionally, as the CH predicts the more
likely influential role of individual difference in interactive task performance as opposed to individual task performance, exploring the mediating role of ability and affective factors in the simple versus complex collaborative writing task performance can provide further significant insights into theory and practice in the field of L2 writing.

This study explored one type of pre-task planning, namely, strategic unguided 10-minute pre-task planning. Future research is warranted to explore the effects of different types of the pre-task planning, including guided strategic planning, rehearsal pre-task planning, rhetorical planning, among others, and the role of ability and affective variables that might mediate the effect of planning on writing production in the performance of both simple and complex tasks. The results of such studies would help guide teachers in matching the tasks with the learners’ instructional needs.

This study could also be replicated and expanded in a number of ways. First, this study focused on upper-intermediate proficiency level adult learners, as studying all proficiency levels in one single research was not feasible. As such, future studies can replicate this study with other levels of proficiency and age groups. Second, this study investigated the modulating role of motivational beliefs and anxiety in the effect of task design features on FL writing production. Future research might replicate this study by adding other ID factors or replacing the current ID factors with new ID factors such as working memory (WM) capacity and language learning aptitude. Third, this study explored the interface between ID, task complexity, and writing production and text quality; that is, cognitive and metacognitive processes that might be triggered by various dimensions of task complexity and interaction.
between task complexity and ID factors were not studied. Therefore, future research is encouraged to explore the cognitive and metacognitive processes involved in writing tasks of varying degrees of cognitive complexity.
References


class compositions. In B. Kroll (Ed.), Second language writing (pp. 140-154).
New York: Cambridge University Press.

performance in L2 writing. IRAL-International Review of Applied Linguistics in
Language Teaching, 45(3), 261-284.

Italian and French as a foreign language. Journal of Second Language Writing,
17(1), 48-60.

writing and speaking. In P. Robinson (Ed.), Second language task complexity:
Researching the cognition hypothesis of language learning and performance
(pp. 90-104). Amsterdam: John 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.

Educational Horizons, 64(1), 36-38.

writing performance: A model testing with structural equation modeling.
Language Learning, 55(2), 335-374.

of academic literacy development. Mahwah, NJ: Lawrence Erlbaum.


Appendices

Appendix A: Writing Tasks

Study 1, Task 1

Imagine you are the government official in charge of allocating (giving) funds of $5,000,000 worth for public projects. You have received three competing projects for public causes: building a new school for the low-income families in the community, buying new buses to improve old and slow public bus transportation, and building a new dam to resolve water shortage.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Thank you for your assistance!

Study 1, Task 2

Imagine you are the government official in charge of allocating (giving) funds of $10,000,000 worth for public projects. You have received six competing projects for public causes: reducing air pollution, creating jobs for the unemployed, building affordable accommodation for the low-income families, providing subsidised healthcare for the low-income families, providing free higher education for the high-achievers, and increasing school budgets.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Thank you for your assistance!
Study 2, Task 1, No Pre-task-planning Time

Imagine you are the government official in charge of allocating (giving) funds of $5,000,000 worth for public projects. You have received three competing projects for public causes: building a new school for the low-income families in the community, buying new buses to improve old and slow public bus transportation, and building a new dam to resolve water shortage.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Please start writing immediately without any planning.

Thank you for your assistance!

Study 2, Task 2, No Pre-task-planning Time

Imagine you are the government official in charge of allocating (giving) funds of $10,000,000 worth for public projects. You have received six competing projects for public causes: reducing air pollution, creating jobs for the unemployed, building affordable accommodation for the low-income families, providing subsidised healthcare for the low-income families, providing free higher education for the high-achievers, and increasing school budgets.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Please start writing immediately without any planning.

Thank you for your assistance!
Study 2, Task 1, with Pre-task-planning Time

Imagine you are the government official in charge of allocating (giving) funds of $5,000,000 worth for public projects. You have received three competing projects for public causes: building a new school for the low-income families in the community, buying new buses to improve old and slow public bus transportation, and building a new dam to resolve water shortage.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the significance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 10 minutes to plan and 35 minutes to write. Please note that you can plan the content, the organisation of your essay, grammar, and vocabulary; there is no limit. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Thank you for your assistance!

Study 2, Task 2, with Pre-task-planning Time

Imagine you are the government official in charge of allocating (giving) funds of $10,000,000 worth for public projects. You have received six competing projects for public causes: reducing air pollution, creating jobs for the unemployed, building affordable accommodation for the low-income families, providing subsidised healthcare for the low-income families, providing free higher education for the high-achievers, and increasing school budgets.

You should allocate funds for all projects. Please prioritise (choose and select) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 10 minutes to plan and 35 minutes to write. Please note that you can plan the content, the organisation of your essay, grammar, and vocabulary; there is no limit. Please write around 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Thank you for your assistance!
Appendix B: Teachers’ Judgment Questionnaire

Please read the following two tasks and explain which of the tasks is cognitively more complex for your upper-intermediate students to write an essay about. Please give your reasons for your choice.

Please answer the question given after the tasks.

Thank you for your assistance!

Task 1

Imagine you are the government official in charge of allocating (giving) funds of $5,000,000 worth for public projects. You have received three competing projects for public causes: building a new school for the low-income families in the community, buying new buses to improve old and slow public bus transportation, and building a new dam to resolve water shortage.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write at least 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.

Task 2

Imagine you are the government official in charge of allocating (giving) funds of $10,000,000 worth for public projects. You have received six competing projects for public causes: reducing air pollution, creating jobs for the unemployed, building affordable accommodation for the low-income families, providing subsidised healthcare for the low-income families, providing free higher education for the high-achievers, and increasing school budgets.

You should allocate funds for all projects. Please prioritise (choose and rank) the projects and allocate the amount of the fund based on your own view of the importance of the projects for the local people. Please provide reasons and give examples when needed to make your choices as convincing as possible for the local people.

You have 35 minutes to write. Please write at least 250 words. There is no upper limit; you can write as many words as you like to make your decision clear and convincing to the local people.
Question:

Which of the above writing tasks is more cognitively complex (difficult) for your upper-intermediate English as a foreign language (EFL) students to write an essay about? Please provide your reasons for your choice as clearly as possible?
Appendix C: Pre-writing Test

You should spend 30 minutes on the following writing task and should write at least 250 words.

Some people believe that using the Internet has caused a lot of problem for young people, while others argue that using the Internet has brought many benefits to young people.

What is your opinion? Use specific reasons and examples to support your opinion.
PARTICIPANT INFORMATION SHEET (PIS)
(School Managers)

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task Production: The Case of Separate and Simultaneous Manipulation of Task Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task Planning

Researcher: Mr Muhammad Rahimi, PhD candidate, School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand

I am a PhD student in the School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand. I am conducting my PhD research project on the role of individual differences and task complexity in foreign language learners’ writing production. This is a significant research project as the findings will contribute to the theory and pedagogy. In other words, the findings may help us in developing more effective writing instruction programmes and also help us to develop more valid writing tasks for task-based writing assessment. As such, your students’ participation as foreign language learners of English is highly valuable for this study. This study will be conducted out of usual class time in two separate days three hours altogether (one hour the first day and around 90 minutes the second day). The students will fill out questionnaires the first day. The second day, they will fill out brief questionnaires and write two English essays. The schools and students’ participation is voluntary and their participation and non-participation will not affect their grades or their relationship with the teacher and/or school in any manner. They can withdraw from the programme at any point without any consequences. Because
all data will be codified and the documents identifying participants’ information will be shredded, the students can also withdraw the traceable data up to three weeks after data collection. The anonymous data will be kept for a maximum of six years, and will be used in writing and publishing PhD thesis and research articles. Students who complete all sections of the study will go for a draw for 20-dollar cash (10 draws) as appreciation of their participation. I will sincerely appreciate your participation in this study by allowing me to conduct this study in your school and by helping me invite upper-intermediate level students and their teachers to participate in the project through email invitation that will be sent by the secretary of your school.

**Contact Details and Approval Wording**

Researcher: Mr Muhammad Rahimi  
*m.rahimi@auckland.ac.nz*  
+64 (0) 9 373 7599  
N - Block. Epsom, Level 4, Room N452, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Supervisor: Professor Lawrence Zhang  
N - Block. Epsom, Level 5, Room 6EN-559, Epsom Campus, 74 Epsom Ave, Auckland 1023, Epsom, New Zealand  
+64 (0) 9 373 7599  
Ext. 48750  
*lj.zhang@auckland.ac.nz*

Head of School: Professor Judy Parr  
+64 (0) 9 923 8998  
*jm.parr@auckland.ac.nz*  
N - Block. Epsom, Level 5, Room 6EN-501, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Chair contact details: “For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142. Telephone: 09 373-7599 extn. 87830/83761 Email: humanethics@auckland.ac.nz.”

**APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.**
PARTICIPANT INFORMATION SHEET (PIS)
(Teachers)

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task Production: The Case of Separate and Simultaneous Manipulation of Task Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task Planning

Researcher: Mr Muhammad Rahimi, PhD candidate, School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand.

I am a PhD student in the School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand. I am conducting my PhD research project on the role of individual differences and task complexity in foreign language learners’ writing production. This is a significant research project as the findings will contribute to the theory and pedagogy. In other words, the findings may help us in developing more effective writing instruction programmes and also help us to develop more valid writing tasks for task-based writing assessment. As such, your participation as a foreign language teacher of English is highly valuable for this study. Your participation is voluntary and anonymous. Your participation and non-participation will not affect your relationship with other teachers, students, or the manager of school in any manner. You will be invited to judge the difficulty of two writing tasks for your upper-intermediate English as a foreign language (EFL) students. You will be asked to provide your judgment and your reasons for your judgment as clearly as possible in writing. You will be asked to return your anonymous judgment essay to a box which will be placed in the school staffroom. The anonymous data will be kept for a maximum of six years and will be used in writing and publishing PhD thesis and research articles. I will sincerely appreciate your participation in this study.
Contact Details and Approval Wording

Researcher: Mr Muhammad Rahimi
m.rahimi@auckland.ac.nz
+64 (0) 9 373 7599
N - Block. Epsom, Level 4, Room N452, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Supervisor: Professor Lawrence Zhang
N - Block. Epsom, Level 5, Room 6EN-559, Epsom Campus, 74 Epsom Ave, Auckland 1023, Epsom, New Zealand
+64 (0) 9 373 7599
Ext. 48750
lj.zhang@auckland.ac.nz

Head of School: Professor Judy Parr
+64 (0) 9 923 8998
jm.parr@auckland.ac.nz
N - Block. Epsom, Level 5, Room 6EN-501, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Chair contact details: “For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142. Telephone: 09 373-7599 extn. 87830/83761 Email: humanethics@auckland.ac.nz.”

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.
PARTICIPANT INFORMATION SHEET (PIS)
(Students)

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task
Production: The Case of Separate and Simultaneous Manipulation of Task
Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task
Planning

Researcher: Mr Muhammad Rahimi, PhD candidate, School of Curriculum and
Pedagogy, Faculty of Education, The University of Auckland, New Zealand.

I am a PhD student in the School of Curriculum and Pedagogy, Faculty of Education,
The University of Auckland, New Zealand. I am conducting my PhD research project
on the role of individual differences and task complexity in foreign language
learners’ writing production. This is a significant research project as the findings will
contribute to the theory and pedagogy. In other words, the findings may help us in
developing more effective writing instruction programmes and also help us to
develop more valid writing tasks for task-based writing assessment. As such, your
participation as foreign language learners of English is highly valuable for this study.
This study will be conducted out of usual class time in two separate days three hours
altogether (one hour the first day and around 90 minutes the second day). You will
fill out questionnaires the first day. The second day you will fill out brief
questionnaires and write two English essays. Your participation is voluntary and your
participation and non-participation will not affect your grades or your relationship
with your teacher and/or school in any manner. You can withdraw from the
programme at any point without any consequences. Because all data will be codified
and the documents identifying participants’ information will be shredded, you can
withdraw the traceable data up to three weeks after data collection. The anonymous
data will be kept for a maximum of six years, and will be used in writing and
publishing PhD thesis and research articles. Students who complete all sections of the study will go for a draw for 20-dollar cash (10 draws) as appreciation of their participation. I will sincerely appreciate your participation in this study.

**Contact Details and Approval Wording**

Researcher: Mr Muhammad Rahimi  
[m.rahimi@auckland.ac.nz](mailto:m.rahimi@auckland.ac.nz)  
+64 (0) 9 373 7599  
N - Block. Epsom, Level 4, Room N452, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Supervisor: Professor Lawrence Zhang  
N - Block. Epsom, Level 5, Room 6EN-559, Epsom Campus, 74 Epsom Ave, Auckland 1023, Epsom, New Zealand  
+64 (0) 9 373 7599  
Ext. 48750  
lj.zhang@auckland.ac.nz

Head of School: Professor Judy Parr  
+64 (0) 9 923 8998  
[jm.parr@auckland.ac.nz](mailto:jm.parr@auckland.ac.nz)  
N - Block. Epsom, Level 5, Room 6EN-501, Epsom Campus, 74 Epsom Ave, Auckland 1023 Epsom, New Zealand

Chair contact details: “For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142. Telephone: 09 373-7599 extn. 87830/83761 Email: humanethics@auckland.ac.nz.”

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.
CONSENT FORM
(School Managers)

THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task Production: The Case of Separate and Simultaneous Manipulation of Task Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task Planning

Researcher: Mr Muhammad Rahimi, Ph.D. student, School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand

Supervisor: Professor Lawrence Zhang, School of Curriculum and Pedagogy, Faculty of Education, the University of Auckland, New Zealand

I have read the Participant Information Sheet. I have understood the nature of the research, and why this language school has been selected. I have had the opportunity to ask questions and have them answered to my satisfaction.

- I agree that you may conduct the above research in this language school.
- I understand that your research will take approximately three hours (two separate days: one hour the first day and around 90 minutes the second day) and will include no audio or videotaping.
- I will provide / will not provide you with classroom to conduct your research.
- I wish / do not wish to receive the summary of findings in a way that does not identify its source.
- I understand that data will be kept for a maximum of 6 years, after which they will be destroyed.
- I understand that the school can withdraw participation at any time. I understand that students can withdraw from participation at any time and can withdraw their traceable data up to three weeks after data collection.
- I confirm that language learners’ participation or non-participation will in no way influence their grades in their course nor their relationship with their teachers or the school.

Name_________________________________ Signature________________________________________

Date ________________
APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.
CONSENT FORM
(Students)

THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task Production: The Case of Separate and Simultaneous Manipulation of Task Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task Planning

Researcher: Mr Muhammad Rahimi, Ph.D. student, School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand
Supervisor: Professor Lawrence Zhang, School of Curriculum and Pedagogy, Faculty of Education, the University of Auckland, New Zealand

I have read the Participant Information Sheet. I have understood the nature of the research, and why I have been selected. I agree to voluntarily participate in this research. I understand that I will have the opportunity to ask questions and have them answered to my satisfaction.

- I agree to take part in this research.
- I understand that I am free to withdraw participation at any time and to withdraw any data traceable to me up to three weeks after data collection.
- I agree to participate in an approximately three-hour research (two separate days- one hour the first day and around 90 minutes the second day) including no audio or videotaping.
- I understand that a research assistant who has signed the confidentiality agreement will score part of the data.
- I understand that the information obtained from my participation in the project will be reported or published in a way that does not identify me as its source.
- I understand that data will be kept for a maximum of 6 years, after which they will be destroyed.
• I understand that the manager of school has given assurance that my participation or non-participation in this research will in no way influence my grades in the course nor my relationship with my teachers or the school.

Name ___________________________Signature _____________________________

Date ________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.
Research Assistant Confidentiality Agreement

Project Title: Argumentative Writing Task Complexity, Task Difficulty, and Task Production: The Case of Separate and Simultaneous Manipulation of Task Complexity along the Degree of Reasoning, the Number of Elements, and Pre-task Planning

Researcher: Mr Muhammad Rahimi, Ph.D. student, School of Curriculum and Pedagogy, Faculty of Education, The University of Auckland, New Zealand.

Supervisor: Professor Lawrence Zhang, School of Curriculum and Pedagogy, Faculty of Education, the University of Auckland, New Zealand.

Research Assistant:

I agree to assist the researcher in all stages of conducting this research including the date collection and scoring students’ essays for the above research project. I understand that the information contained within them is confidential and must not be disclosed to, or discussed with, anyone other than the researcher and his supervisor(s).

Name: _____________________________

Signature: __________________________

Date: _____________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 06-Nov-2014 FOR (3) YEARS REFERENCE NUMBER 013290.