COGNITIVE TASK COMPLEXITY, FOREIGN LANGUAGE ANXIETY AND L2 PERFORMANCE IN SPANISH: A TASK-BASED LANGUAGE TEACHING PERSPECTIVE

A Dissertation
submitted to the Faculty of the
Graduate School of Arts Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Spanish and Portuguese

By

Ángela Donate, M.S.

Washington, DC
May 16, 2018
COGNITIVE TASK COMPLEXITY, FOREIGN LANGUAGE ANXIETY AND L2 PERFORMANCE IN SPANISH: A TASK-BASED LANGUAGE TEACHING PERSPECTIVE

Ángela Donate, M.S.

Dissertation Advisor: Ronald P. Leow, Ph.D.

ABSTRACT

Although cognitive psychology literature (e.g., Derakshan & Eysenck, 2009) has demonstrated the detrimental effects anxiety has on cognitive processes, this relationship has barely been investigated in the SLA field (e.g., MacIntyre & Gardner, 1994). In addition, the vast majority of SLA and TBLT research examining foreign language anxiety (FLA) have consistently focused on the negative outcomes of FLA, treating it as single powerful force to be simply eliminated, therefore, ignoring its facilitative role (e.g., Scovel, 1978), its dynamicity and fluctuation over time during learners’ task performance (e.g., Baralt & Gurzynski-Weiss, 2011), and its three interconnected facets within the language learner—affective, cognitive, and behavioral (Zeidner & Matthews, 2011).

Employing mixed-methods, this dissertation attempted to offer a more inclusive and complete understanding of the role of language anxiety in L2 oral tasks that vary in the cognitive demands imposed on the learner. Following Sasayama (2016), fifty-one L1 English low proficient learners of Spanish performed two oral narrative tasks manipulated via number of elements. State anxiety was measured at two points during task performance, halfway and immediately after task completion (Baralt & Gurzynski-Weiss, 2011). Subjective learners’ perspectives were elicited in relation to their emotional states during oral task performance. L2 task performance was assessed using complexity, accuracy and fluency (CAF) measures to examine how state anxiety was related to task performance.
Quantitative results revealed levels of perceived state anxiety increased meaningfully with greater cognitive demands. As predicted, learners’ levels of language anxiety fluctuated over time, showing evidence to support the dynamic nature of the construct as an affective factor. Qualitative data showed a cyclical two-way relationship of the three-way interface of Language Anxiety—Cognition—Task Performance, and further suggests a tendency where the facilitative and detrimental role of language anxiety work simultaneously during task performance. In relation to existing task taxonomies, findings revealed that for low proficient L2 learners, linguistic demands, conceptual information and task mode seem to be assessed as major predictors of state anxiety during task performance. Important pedagogical implications are discussed in relation to the three-way interface of Language Anxiety—Cognition—Task Performance in the L2.
DEDICATION

A Drew, mi gran amor, a mis abuelos Pochola y Antolín, a mis padres, Mercedes e Ignacio y a mi hermano Nacho por haber sido siempre y ser los grandes pilares de mi vida. Por su dedicación y cariño incondicional.
ACKNOWLEDGEMENTS

After an intensive period of learning and conducting research, I would like to reflect on the people who have contributed in one way or another to make this dissertation happen. Reaching this point would have never been possible without their support.

First and foremost, I am deeply grateful to my advisor, Dr. Ronald Leow, who has guided me from the first day of my doctoral program at Georgetown. He taught me how to become a true scholar and researcher, and I will be forever indebted to him for his continuous support throughout this enriching and challenging journey. Among the most valuable lessons I learned from Dr. Leow have been reading research critically, looking at the ‘big picture’ and writing academically. Dear Ron—I value how you genuinely care about your mentees, dedicated your time generously all these years, and of course, how you pushed me to my limits and to go one step further. I cannot thank you enough!

Second, to my committee members, Cristina Sanz and Anastasia Drackert for being incredible role models and for their insightful feedback. Cristina – thank you so much for all your valuable guidance and advice on so many occasions, and especially for opening a door to your office when I really needed to hear “you can do it!” I have learned so much from you. Anastasia – we started this academic journey together, from drawing syntax trees to discussing SLA projects over coffee, and I am grateful you joined my dissertation committee. Your continuous mentorship has shaped my Georgetown experience, and you have had an enormous influence my academic preparedness.

Third, I want to extend my most sincere gratitude to Dr. Lourdes Ortega, who helped me discover abilities as a researcher I did not even know I had. Thank you for
inspiring me, helping me develop academically and for your priceless lessons on academic writing.

Thank you to my other brilliant Georgetown professors, Dr. Héctor Campos, Dr. Alfonso Morales-Front, Dr. Tom Walsh, Dr. Michael Ferreira, Dr. John Norris, Dr. John Davis, Dr. Jennifer Sclafani, and Dr. Luke Plonsky for providing such a strong foundation in my academic training.

Next, my deep appreciation goes to Annie Ornelles and Jeong Eva Mun for all their time and generosity in assisting with data collection and coding. They enabled accurate and reliable analyses of the data and saved a great deal of time. I am thankful for Yasser Teimouri and to Todd MacKay for kindly helping me so much with statistics and for their guidance to understand questionnaire data. Thank you to Dr. Rusan Chen for aiding me with my statistical analyses. I would also like to express my gratitude to Dr. Norris for his invaluable lessons on applying statistical procedures in an ethical way, interpreting results accurately and situating them in meaningful findings. Another big thank you goes to Dr. Peter Janssens for his assistance reserving the laboratory for data collection.

One of the major reasons I ended up pursuing a Ph.D. was the inspiration from my very dear friend, Dr. Berta Parra de Chopite, who I had the privilege to meet during my time at Wabash College in Indiana. She guided me throughout the entire application process and became both my American mother and one of the most trustworthy friends I ever had. Another key reason for my desire to follow an academic career was the support of my professors at the University of Oviedo, namely, Dr. Marta Mateo, Dr. Javier Vallina, Dr. Carmen Escobedo de Tapia, Dr. María José Alvarez Faedo, and Dr. Cristina Valdés. They introduced me to the academic world, raised my interest in the study of linguistics
and constituted the underpinnings of my undying passion for foreign languages. A special thank you goes to Dr. Mateo, who connected me to academic opportunities abroad, and to Dr. Vallina for his continuous encouragement and going above and beyond to help me achieve my goals. I am deeply indebted to each of you and admire you for your excellence as educators and kindness as human beings.

To mi adorada Celia Zamora – what I appreciate most about you is your generous heart. I am always amazed by your willingness and dedication to help your friends. You know I could have never reached this point without you. To Allison Caras – I love and admire your spirit, your positive attitude and always with a big smile. I greatly appreciate your encouragement and being a friend I can always count on. I could have never wished for better partners than the two of you.

Thank you to Inés Corujo, Pablo Camus, Ariel Zach, Fátima Montero and Mercedes Ontoria Peña for being such wonderful friends and colleagues. You know how much you mean to me. I cannot be more grateful for your friendship and the emotional support, laughter, beers at Epicurean, gossip, hugs, tears and much more. Thank you for sharing so much with me. You all have made me a better scholar, a better person and have by far made the best part of my time at GU.

A special thank you goes to my twin (and ‘tipping point connector’), Julia Goetze. What can I say? I could not have finished this dissertation without you. As my advisor, Ron, would say, my “baby” (this dissertation) was born in the corridors of the ICC, right in between the German and the Spanish Departments! You were with me every single step of the way (we even defended one day apart!), in the best and the worst moments. We have understood each other so well, our common struggles and especially our dissertation topics
when no one else did. Why study emotions in a foreign language? What for? Thank you for believing in me unconditionally when I did not. I will always cherish our memories together whether in laughter or in tears. Thank you for inspiring me, for your eternal support, for all the hours coding my data, for being one of the greatest friends, and for sharing with me more than I could imagine.

I would also like to thank many more wonderful Hoya friends and colleagues who have shared this exciting experience with me, including Lara Bryfonski, Brandon Tullock, Tim McCormick, Katie Vadella, Marisa Filgueras-Gomez, Colleen Moorman, Young-A Son, Silvia Marijuan, Chrissy Bistline-Bonilla, Gabriela DeRobles, Janire Zabildea, Alexandra Martin, Megan Driver, Nick Pandža, Sergio Adrada-Rafael, Hortensia Gutiérrez, and Natalia Curto. The same goes for my fantastic comrades from the literature department, including Yoel Castillo Botello, Bohumira Smidakova, Alfredo Poggi, Felipe Toro, Jovana Zujevic, Gabriel Villarroel, Porter O’Neill, Xabier Fole Varela and Natalia Pavón. You have been like family to me and enriched my time at Georgetown in a very special way.

I want to conclude by dedicating this dissertation to my dear family – above all, to my husband, Drew Brookie, for his love, infinite support and patience. You have been my rock all this time, and this ultimate achievement would not have been possible without you in my life. To my parents, Ignacio and Mercedes, and my brother, Nacho, for their motivation, love and continuous words of encouragement. I could not possibly think of more dedicated family than you. Thank you to all my Donate and Velasco extended family back home, in particular, to my dear cousins María, Alejandra y Teresa, for being always like sisters to me. The same goes to my dear in-laws, Cyndy, Dean, Paige and Graham for
their support. To my grandma, Amparo, my biggest fan. I know how proud you would feel if you could have witnessed this achievement.

Last, but not least, I dedicate this project to my dear grandparents, Pochola y Antolín, the pillars of my family, and, undoubtedly, the most generous and admirable human beings I have ever known. Gracias por habérme dado todo, por enseñarme las más valiosas lecciones de vida y por cuidarnos a todos tanto y tan bien. Desde donde estéis, os adoro y os llevo siempre contigo.
# TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ................................................................. 1

1.1 Introduction: Origins and Emergence of Task-Based Language Teaching ........ 1
1.2 From Task-Based Language Teaching to Task-Based Language Learning .......... 5
1.3 The Use of Tasks in Second Language Acquisition Research ...................... 6
1.4 Problem Statement ........................................................................ 7

CHAPTER 2: COGNITIVE TASK COMPLEXITY .................................... 13

2.1 The Construct of Cognitive Task Complexity ...................................... 13
2.2 Robinson’s Cognition Hypothesis and Skehan’s Limited Attentional Capacity Model ........................................................................ 14
2.3 Cognitive Task Complexity Research ............................................... 19
2.4 Previous Research Operationalizing Task Complexity Via the Number of Elements on the Resource-Directing Dimension in Robinson’s Cognition Hypothesis ....... 23
2.5 Current Limitations in the Operationalization of Cognitive Task Complexity ...... 31
2.6 Cognitive Load Theory and Mental Effort in Educational Research ............. 33
2.7 Task-Based Language Teaching Research and Independent Cognitive Complexity Measures ........................................................................ 35
2.7.1 Task Difficulty and Mental Effort Measures ....................................... 35
2.7.2 Introspective Methods ..................................................................... 36
2.7.3 Dual-Task Methodology ..................................................................... 37
2.7.4 Time Estimation ............................................................................. 37
2.8 Task-Based Language Teaching Research on Cognitive Load Measures ........ 39
2.9 Future Directions in Task-Based Language Teaching .................................. 46

CHAPTER 3: FOREIGN LANGUAGE ANXIETY (FLA) AND L2 TASK PERFORMANCE ........................................................................... 49

3.1 Conceptualization of Anxiety as a Psychological Construct: Its Origins .......... 49
3.2 Theoretical Distinctions of Anxiety as Construct: Trait vs. State Anxiety ....... 50
5.1.2 Testing the Effect of Task Order on the Cognitive Load Measures ............ 153
5.1.3 Qualitative Results for Task Difficulty and Mental Effort ...................... 155
  5.1.3.1 Task Difficulty ............................................................................. 156
  5.1.3.2 Mental Effort ............................................................................. 161
5.1.4 Summary of Results for Research Question 1 ...................................... 164

5.2 Results for State Anxiety ........................................................................ 165
  5.2.1 Reliability Analyses for the Language Anxiety Questionnaires .......... 165
  5.2.2 Quantitative Results for Perceived State Anxiety: Simple and Complex Tasks in Time 1 and Time 2 ................................................................. 167
  5.2.3 Results on the Fluctuation of State Anxiety Over Time (From Time 1 to Time 2) and its Dynamicity on the Simple Task .............................................. 167
  5.2.4 Results on the Fluctuation of State Anxiety Over Time (From Time 1 to Time 2) and its Dynamicity on the Complex Task ........................................ 172
  5.2.5 On the Relationship between Cognitive Task Complexity and State Anxiety at Time 1 and Time 2 ..................................................................... 173
    5.2.5.1 Inferential Analyses for State Anxiety at Time 1 and Time 2 ........ 176
  5.2.6 Testing the Effect of Order of Tasks on State Anxiety at Time 1 and Time 2 ........................................................................................................ 183
  5.2.7 Qualitative Data: Open-Ended Questions for State Anxiety on the Simple and Complex Task ................................................................. 188
    5.2.7.1 Emotions ....................................................................................... 188
    5.2.7.2 Facilitative Versus Debilitative Language Anxiety ......................... 189
    5.2.7.3 Task-Internal Features .................................................................. 191
    5.2.7.4 Task-Linguistic Demands ............................................................... 193
      5.2.7.4.1 Vocabulary ............................................................................. 193
      5.2.7.4.2 Grammar ............................................................................. 193
  5.2.8 Relationship Between Perceived State Anxiety and Classroom Anxiety in the Spanish Classroom ........................................................................... 196
5.2.9 Summary of Results for Research Question 2 and Research Question 3 ..... 197

5.3 Results on the Relationship between Language Anxiety, Cognition and Task Performance in Second Language Spanish When Cognitive Task Complexity is Manipulated: The Three-Way Interface of Language Anxiety — Cognition — Second Language Task Performance .................................................. 200

5.3.1 On the Relationship between Language State Anxiety and Cognitive Processes During Oral Task Performance when Cognitive Load is Manipulated (Research Question 4a) ........................................................................................................ 200

5.3.1.1 Short-Term Memory/ Working Memory ........................................ 200

5.3.1.2 General Thought Process ............................................................. 201

5.3.1.3 Thought and Idea Organization .................................................. 202

5.3.1.4 Focus ....................................................................................... 202

5.3.2 On the Three-Way Interface Language Anxiety — Cognition — Task Performance in Second Language Tasks ........................................................................................................ 203

5.3.2.1 Relationship ‘Language Anxiety — Cognition — Task Performance’ in L2 Spanish ......................................................................................................................... 204

5.3.3 On the Relationship between State Anxiety and Second Language Performance—as Perceived by Learners—When Cognitive Task Complexity is Manipulated (Research Question 4b) ........................................................................................................ 207

5.3.3.1 Learners’ Perceptions on the Main Linguistic Areas Affected by State Anxiety................................................................................................................................. 208

5.3.3.1.1 Debilitative and Facilitative Anxiety: Learners’ Perceptions on its Effects on Second Language Task Performance ................................................................................ 208

5.3.3.2 Language Anxiety Coping Strategies ........................................ 211

5.3.3.2.1 Group 1: ‘Avoidance Strategies’ ................................................ 211

5.3.3.2.2 Group 2: ‘Other Emotional Compensatory Strategies’ ............ 212

5.3.4 Summary of Results for Research Question 4 .................................. 214

5.4 Results for Second Language Performance: Complexity, Accuracy and Fluency Measures across the Two Cognitive Task Complexity Levels ........................................................................................................ 216

5.4.1 Syntactic Complexity Measures ...................................................... 216

5.4.2 Accuracy Measure ........................................................................ 221
5.4.3 Fluency Measures ................................................................. 225
5.4.4 Summary of Results for Research Question 5 .......................... 236

CHAPTER 6: DISCUSSION ............................................................. 237

6.1 Research Question 1. Validating the Intended Relationship between Task Complexity and Cognitive Demands Imposed by the Two Tasks ................................................................. 237
6.2 Research Question 2a. Relationship between Cognitive Task Complexity and Levels of State Anxiety as Experienced during Task Performance on the Two Levels of Cognitive Complexity ........................................................................ 247
   6.2.1 Major Factors Influencing Language State Anxiety during Oral Task Performance ........................................................................ 250
6.3 Research Question 2b. On the Relationship between State Anxiety and Foreign Language Classroom Anxiety ............................................................................................................. 254
6.4 Research Question 3. On the Dynamicity of Language Anxiety Over Time ...... 257
6.5 Research Question 4. On the Three-Way Interface Language Anxiety—Cognition—Second Language Performance: The Three Facets on Foreign Language Anxiety from a Task-Based Language Teaching Perspective ........................................................................... 259
   6.5.1 Debilitative and Facilitative Anxiety ........................................ 262
6.6 Research Question 5. On the Relationship between Cognitive Task Complexity and Second Language Performance Measures ............................................................................................. 263

CHAPTER 7: LIMITATIONS, FUTURE DIRECTIONS AND PEDAGOGICAL IMPLICATIONS .............................................................. 266

7.1 Limitations .................................................................................. 266
7.2. Future Directions ........................................................................ 268
7.3 Implications for Teaching and the Application Positive Psychology: Why Does It All Matter? ................................................................................................................................. 272
   7.3.1 Pedagogical Considerations on Task Design for Low Proficiency Levels ................................................................. 273
   7.3.2 Positive Psychology to Inform Second Language Classroom Pedagogic Decisions: Raising Awareness on How Language Anxiety Actually Operates in the Second and Foreign Language ............................................................................. 274
   7.3.3 Implications for Assessing Learners’ Complexity, Accuracy and Fluency ................................. 279

CHAPTER 8: CONCLUSIONS ................................................................ 280
APPENDIX: MATERIALS .................................................................................................................. 285

Appendix A: Language Background Questionnaire ........................................................................ 285
Appendix B: Task Instructions ........................................................................................................ 294
Appendix C: State Anxiety Scale Mid-Task .................................................................................... 296
Appendix D: Post-Task Anxiety Scale ............................................................................................ 298
Appendix E: Self-Assessed Cognitive Load Measures .................................................................... 302
Appendix F: Cognitively Simple Task ............................................................................................ 304
Appendix G: Cognitively Complex Task ........................................................................................ 305
Appendix H: Retrospective Time Estimation Measure ................................................................... 306
Appendix I: Foreign Language Classroom Anxiety Scale .............................................................. 307
REFERENCES .............................................................................................................................. 312
LIST OF FIGURES

Figure 1: Ng and King’s (2018) Multi-dynamic view proposed for researching teacher emotion (cited in Mercer & Kostoulas, 2018, p. 145) ................................................................. 65

Figure 2: Donate’s (2018) adapted version of King and Ng’s model for the L2 learner: Interconnectedness between Language Anxiety — Cognition — L2 Performance. ........ 66

Figure 3: Spielberg’s model of state anxiety (1966) depicting transactional trait - state conceptions of anxiety (cited in Zeidner & Matthews, 2011, p. 80)......................... 68

Figure 4: Boxplots for task difficulty before outliers were removed based on the two cognitive task complexity levels. ................................................................. 139

Figure 5: Histogram for task difficulty for the cognitively simple task............... 140

Figure 6: Histogram for task difficulty for the cognitively complex task .......... 141

Figure 7: Histogram for mental effort for the cognitively simple task ............. 143

Figure 8: Histogram for mental effort for the cognitively complex task .......... 144

Figure 9: Self-assessed ratings for task difficulty based on the two cognitive complexity levels................................................................. 146

Figure 10: Self-assessed ratings for mental effort based on the two cognitive complexity levels................................................................. 147

Figure 11: Time estimation distribution of scores before withdrawing outliers based on the two cognitive complexity levels ................................................................. 148

Figure 12: Histogram after outliers were excluded for the cognitively simple task on the time estimation measure (secs). ................................................................. 149

Figure 13: Histogram after outliers were excluded for the cognitively complex task .... 150

Figure 14: Error bars for time estimation based on the two cognitive complexity levels ................................................................. 152

Figure 15: Histogram for levels of state anxiety for the simple task at Time 1. .... 169

Figure 16: Histogram for levels of state anxiety for the simple task at Time 2. .... 171

Figure 17: Histogram for levels of state anxiety for the complex task at Time 1. .... 172

Figure 18: Histogram for levels of state anxiety for the complex task at Time 2. .... 176

Figure 19: Error bars with confidence intervals for state anxiety at Time 1 based on the two cognitive complexity levels. ................................................................. 178
Figure 20: Error bars with confidence intervals for state anxiety at Time 2 based on the two cognitive complexity levels. ................................................................. 179

Figure 21: Error bars for the simple task exploring state anxiety over time (Time 1 versus Time 2) ............................................................................ 180

Figure 22: Error bars for the complex task exploring state anxiety over time (Time 1 versus Time 2) ............................................................................ 181

Figure 23: Profile plot for the interaction between cognitive task complexity (simple versus complex) and Time (Time 1 versus Time 2) on state anxiety. ...................... 182

Figure 24: Profile plots for the interaction between the order of tasks and cognitive task complexity (simple versus complex) on language state anxiety at Time 1 ...................... 185

Figure 25: Profile plots for the interaction between the order of tasks and cognitive task complexity (simple versus complex) on language state anxiety at Time 2 ...................... 187

Figure 26: Confidence intervals for the global syntactic complexity measure—mean length of AS-units (MLASU) — for the simple and the complex task. ...................... 218

Figure 27: Confidence intervals for the phrasal syntactic complexity measure—mean length of clause (MLC) — for the simple and the complex task. ...................... 219

Figure 28: Confidence intervals for the subordination measure of syntactic complexity—clause per AS unit (C/AS)—for the simple and the complex task. ...................... 220

Figure 29: Histogram for the accuracy measure—weighted clause ratio (WCR)—on the cognitively simple task. ................................................................. 222

Figure 30: Histogram for the accuracy measure—weighted clause ratio (WCR)—for the cognitively complex task. ................................................................. 222

Figure 31: Error bars for the accuracy measure—weighted clause ratio (WCR)—based on the simple and the complex task. ................................................................. 225

Figure 32: Histogram for total number of words (TNW) for the simple task. ............. 226

Figure 33: Histogram for total number of words (TNW) for the complex task. ........... 227

Figure 34: Histogram for total number of syllables (TNS) for the simple task. .......... 227

Figure 35: Histogram for total number of syllables (TNS) for the complex task. .......... 228

Figure 36: Confidence intervals for total number of words (TNW) based on the two complexity levels. ................................................................. 229

Figure 37: Confidence intervals for total number of syllables (TNS) based on the two complexity levels. ................................................................. 230
Figure 38: Histogram for words per minute (W/min) for the simple task. .................. 231
Figure 39: Histogram for words per minute (W/min) for the complex task. .............. 231
Figure 40: Histogram for syllables per second (Syl./sec) for the simple task. .......... 232
Figure 41: Histogram for syllables per second (Syl./sec) for the complex task. ....... 232
Figure 42: Confidence intervals for words per minute (W/min) for the simple and the complex task. ........................................................................................................... 234
Figure 43: Confidence intervals for syllables per second (Syl./sec) for the simple and the complex task. ........................................................................................................... 235
LIST OF TABLES

Table 1: Most common operationalizations of task complexity................................. 20

Table 2: Design of main tasks.................................................................................... 90

Table 3: Superordinate and ordinate categories for the qualitative content analysis with examples extracted from the data................................................................. 108

Table 4: Overview of the CAF measures for L2 task performance............................. 121

Table 5: Clause categorization for weighted clause ratio........................................ 126

Table 6: Scoring system and error weighting for weighted clause ratio..................... 127

Table 7: Example of categorization and scoring......................................................... 128

Table 8: Overview of analytical procedures.............................................................. 136

Table 9: Descriptive statistics for task difficulty for the simple and the complex task. 142

Table 10: Descriptive statistics for mental effort on the simple and the complex task. 145

Table 11: Descriptive statistics for time estimation on the simple and the complex task. 151

Table 12: Descriptive statistics for the order of tasks on task difficulty for the two levels of task complexity................................................................. 154

Table 13: Descriptive statistics for the order of tasks on mental effort for the two levels of task complexity................................................................. 155

Table 14: Descriptive statistics for state anxiety for the simple task at Time 1............ 168

Table 15: Descriptive statistics for state anxiety for the simple task at Time 2............ 170

Table 16: Descriptive statistics for state anxiety for the complex task at Time 1........ 174

Table 17: Descriptive statistics for state anxiety for the complex task at Time 2........ 175

Table 18: Descriptive statistics for the order of tasks on state anxiety for the two levels of task complexity at Time 1................................................................. 184

Table 19: Descriptive statistics for the order of tasks on state anxiety for the two levels of task complexity at Time 2................................................................. 186

Table 20: Descriptive statistics for syntactic complexity for the simple and the complex tasks........................................................................................................... 218
Table 21: Mean number of clauses per error levels, weighting error clauses, raw total score, and weighted clause ratio (WCR) for the two levels of cognitive task complexity. .................................................................................................................. 223

Table 22: Descriptive statistics for the accuracy measure—weighted clause ratio (WCR)—based on the simple and the complex task. ........................................................................................................ 224

Table 23: Descriptive statistics for the productivity measures for the simple and the complex task. ........................................................................................................................................ 229

Table 24: Descriptive statistics for the speech rate measures for the simple and the complex task. ........................................................................................................................................ 233
CHAPTER 1: INTRODUCTION

1.1 Introduction: Origins and Emergence of Task-Based Language Teaching

Task-Based Language Teaching (TBLT) has been defined as an educational framework with the purpose of applying both theory and practice to second and foreign language teaching in an instructional setting (Van den Branden, Bygate & Norris, 2009). In the 1980s, the field of TBLT was born out of a paradigm shift from teacher-centered pedagogical approaches from the 50s and 60s that focused on behaviorism as the model for language education.

Over the past several decades, scholars, language instructors and curriculum developers have disagreed with respect to the pedagogic principles that should guide second and foreign language learning (Richards & Rodgers, 2014). Educational researchers and applied linguists proposed a wide variety of models and approaches (from the earlier Audiolingual or the Natural Approach to the latter Communicative Approach). Van den Branden et al. (2009) discussed the existence of three dichotomies used to explain the main differences between these early behaviorist models and the more recent communicative approaches. In broad terms, these distinctions in language learning distinguished between “discrete versus holistic learning, teacher-dominated versus learner-driven education, and form-focused versus communication-based instruction” (p. 2-3).

In the first dichotomy, language is either viewed as a complex system further divided into small and simple units, isolated grammar rules and words that must be learned in discrete bits (discrete learning); or as discourse in use, in which learners develop their L2 skills and knowledge in an inductive way, by the actual use of the language to fulfill particular functions in specific communicative situations (holistic learning). The second
dichotomy refers to the role of the teacher or the learner as the primary figure who promotes L2 learning. In the former case, the teacher is treated as the protagonist of the learning process, providing constant comprehensible input, speaking and controlling learners’ turn taking during most of the class period. In the latter, the student takes the initiative, deciding on the topic and the particular language constructions by interacting with his/her peers. The instructor in this approach merely guides learners by promoting peer interaction, cooperation and language use. Finally, the third dichotomy establishes that language can be learned “paying attention to form, or communication (or a combination of both)” (p. 3).

According to the authors, while form-focused instruction emphasizes accuracy and complexity of the language as the main aspect of language learning, communication-instruction is based on meaning as the basic element that creates mutual understanding.

The early language learning models that had been focusing on discrete learning, teacher-dominated, and form-focused instruction as main theoretical tenets (Van den Branden et al., 2009) began to be observed as shortcomings leading language teaching researchers (Bygate, 2008; Norris, 2009, Long, 2015) and other prominent educational science scholars (Dewey, 1938; Vygotsky, 1986; Freinet, 1993) to raise their concerns and propose new options for language teaching and learning.

One of the key assumptions of TBLT, as a theoretical and pedagogic framework, is that the purpose of second and foreign language learning should be to “develop the ability to use the target language” (Samuda & Bygate, 2008, p. 7). The idea of second language acquisition (SLA) connected to the real-life use of language started to be a matter of concern in the 1980s, when the term was coined by SLA researchers and educational scholars (e.g., Long, 1985; Prabhu, 1987). In order to reflect the essence of TBLT,
Researchers have pointed out that language is not only practiced with the aim of learning, but rather that L2 learning occurs by using the language (Norris, 2009). Consequently, the emergence of TBLT was rooted in the idea of creating educational activities that captured and reflected the same type of social interaction present in real life situations, transferring it to the context of the language classroom setting (Ellis, 2003).

Supporting this view of language teaching, an increasing number of scholars in the SLA field (Halliday, 1975; Wells, 1981) conducted empirical studies on first language acquisition processes, supporting the meaningful use of language in the second language (as opposed to the more explicit form-focused instruction). In addition, early classroom interaction literature (Barnes, Britton & Torbe, 1986; Wells, 1987) and educational research (Dewey, 1938; Vygotsky, 1978) advocated for moving to a more holistic, learner-driven, and communicative view of language teaching and learning in order to solve the limitations posed by traditional teaching methods. Richards and Rodgers (2014) noted that the new SLA trends, in addition to the theoretical insights coming from outside of the field (mainly educational sciences), gave rise to a new line of thinking that derived into what is commonly known today as Communicative Language Teaching (CLT).

Interestingly, CLT established the basic principles for TBLT, catering for functional and meaningful language use as the heart of the model. Proponents of CLT (Brumfit, 1984; Widdowson, 1978) argued that communication should be the primary purpose of language teaching and learning, as well the promotion of the learner’s autonomy through intensive practice of the oral production and interaction skills (Samuda & Bygate, 2008).
Although in theory this paradigm change sets the theoretical basis for a more functional language teaching conceptualization and a more active role for the learner, in practice CLT was translated mainly into a continuation of the traditional methodologies with a strong focus on practice of oral production and interaction abilities in the classroom. As a result, instead of moving away from discrete (form-focused) teacher dominated instruction, CLT failed to incorporate the intended teaching principles as part of their methodologies and “still maintained the structural knowledge-oriented framework” (Van den Branden et al., 2009, p. 5) in most language programs’ syllabi and curricula.

Although TBLT followed the fundamental theoretical underpinnings established by the CLT approach, it also emerged as a logical step to solve its methodological limitations and offered a solution to its partial account of communication (Richards & Rodgers, 2014). To fill in the gap between theory and actual classroom practice, as Van den Branden et al. (2009) highlighted, task-based language teaching originated as a framework of second language education that (1) attempted to develop complex functional abilities through a “holistic, meaning-focused and learner driven” approach (p. 5) and (2) proposed “task” as the basic unit of the model that connects “educational goals, pedagogic activity and assessment,” (p. 5) described as the three angles of the educational triangle.

One of the fundamental theoretical underpinning that distinguishes TBLT from previous educational models is that the syllabus (what is to be taught) and the methodology (how it is taught) are blended under the same unit of analysis, “the task” (Long, 1985). Due to the central role of “the task” as the core of the model, several definitions of this concept have been proposed both in the field of SLA and by TBLT literature. One of the most influential and cited definitions has been proposed by Long (1985):
“The hundred-and-one things people do—identifiable as bounded recurring activities that people engage in—in order to reach real world (and very often non-linguistic) ends, and for which they need to use language” (p. 89).

The example above illustrates the usage-based nature of TBLT as a framework whose main principles are based on imitating meaningful real-world in order to enhance the effectiveness of language learning (Ellis, 2000). In addition to its meaningful and functional dimension, some of the most influential characteristics attributed to the concept of language learning task have been: “differentiated sequenceable problem-posing activities” (Candlin, 1987:10); goal-directed activities (Swales, 1990); and activities that raise awareness of the functional use of language (Norris, 2009).

1.2 From Task-Based Language Teaching to Task-Based Language Learning

According to Norris (2009), for an adequate understanding of TBLT, it is important to be aware of its multiple components (beyond mere task design), which have turned it into a multifaceted and interdisciplinary pedagogical approach. Consequently, TBLT brings together various interconnected aspects around language program education: learner needs analysis, syllabus design, curriculum development, assessment, program evaluation, teaching, and pedagogic materials and instruction (Norris, 2009).

Connected to its interdisciplinary nature, the construct of “task” has been placed at the center of curricular planning (Robinson, 2011) being used by researchers, language program administrators, syllabus designers, curriculum developers, language instructors, test developers, teacher trainers and assessment centers (Van den Branden et al., 2009). Among the various TBLT areas of study, the current dissertation is concerned with task research from a pedagogic perspective: materials development and task design.
Over the past three decades, drawing from TBLT underpinnings, a significant amount of literature in SLA conducted empirical research on Task-Based Language Learning (TBLL), examining task design and the characteristics of pedagogic tasks that lead to positive effects in learners’ L2 linguistic outcomes (Robinson, 2011). In particular, a number of previous influential studies on TBLL agreed that task-work promotes successful L2 acquisition processes (Gass, Mackey, Alvarez-Torres, & Fernandez-García, 1999; Skehan & Foster, 1999). The section below presents a review of early SLA literature with a focus on the use of L2 tasks on TBLL.

1.3 The Use of Tasks in Second Language Acquisition Research

A vast majority of previous studies has supported the facilitative role of the use of second language (L2) tasks on SLA (Robinson, 2011). A prominent “cognitive-interactionist” strand of research into TBLL has demonstrated the beneficial effects of L2 tasks on several areas of instructed SLA (understood as designed and intentional classroom language acquisition that occurs in instructional settings) (Bygate, Norris & Van den Branden, 2015; Robinson, 2011).

One of the SLA areas that has received considerable attention is task-based interaction research (Crookes & Gass, 1993; Doughty & Pica, 1986; Gass & Varonis, 1994; Long, 1983, 1996; Long & Porter, 1985; Mackey, 1999; Pica & Doughty, 1985; Pica, Young, & Doughty, 1987; Swain, 1993, 1995; Swain & Lapkin, 1995). Specifically, this strand of research has informed instructed SLA about the benefits of L2 tasks in providing a context for a) comprehending and negotiating the meaning of language provided in task input, or used by a partner performing the same task with the provision of feedback (Long, 1996; Pica, 1994; Gass & Mackey, 2007; Gass, Mackey & Ross-Feldman, 2011; Mackey
& Gass, 2006), and b) facilitating learners’ *noticing the gap* in their interlanguage (Swain, 1993; Swain & Lapkin, 1995).

In addition to the interactionist strand, more recently TBLT scholars (Bygate et al., 2015; Long, 2015; Robinson, 2011) have discussed the facilitative role of L2 tasks in raising the learner’s awareness and realization of the functional use of language during task performance.

Furthermore, it is necessary to point out two major strands of research that have contributed immensely to TBLT, which shed light on other factors that influence L2 learners’ linguistic outcomes in task performance: the task-based cognitive literature (Baralt, 2013; Bygate, Skehan & Swain, 2013; Gilabert, 2007a, 2007b; Kim, 2009; Robinson, 2007) concerned with the cognitive demands of tasks, and the motivation literature (Dörnyei, 2005; Dörnyei & Kormos, 2000), informing TBLL into what and how specific task features affect L2 learners within a particular communicative context.

Importantly, the cognitive strand of research has been commonly referred to as “task complexity” by the majority of SLA and TBLT literature. Essentially, the studies cited above (among others) have examined how the cognitive processes and demands (such as attention, short-term and long-term memory) imposed by tasks affect learners’ task performance and L2 production (Robinson, 2011).

1.4 Problem Statement

Up to date, the most widely investigated strand of research within the TBLT literature during the past few decades has been concerned with how to sequence L2 tasks in a task-based syllabus according to the cognitive demands required to perform the tasks. A considerable amount of studies on cognitive task complexity (Gilabert, 2009; Kuiken &
Vedder, 2008; Kormos & Trebits, 2011; Révész, 2011; Robinson, 2001, 2007, 2011; Torres, 2013) have examined the effects of that increasing the cognitive demands of tasks has on learners’ language production with the purpose of promoting language development (Branden, Bygate & Norris; Long, 2015; Robinson, 2009).

One noteworthy challenge presented to TBLT researchers has been that of understanding how and to what extent the role of affective factors has on learners’ performance. Specifically, the role of language anxiety—considered as a dynamic and temporal reaction to an L2 task—on L2 task performance when cognitive task complexity is manipulated remains unknown. To date, from a TBLT perspective, only Baralt and Gurzynski-Weiss (2011) have attempted to address—at least partially—the methodological and empirical limitations presented by the research that has investigated the dynamicity of language anxiety. Overall, previous studies on task complexity (e.g. Révész, 2009; Robinson, 2007; Tracy-Ventura, 2011) that have explored to a certain extent the effects of language anxiety have been regarded as problematic in the way language anxiety has been conceptualized and operationalized in research designs. In this respect, most SLA and TBLT studies have failed to capture the dynamic nature of anxiety by employing general measures of anxiety that gauge foreign language classroom anxiety instead, such as, the Horwitz, Horwitz et Cope (1986) Foreign Language Classroom Anxiety Scale (FLCAS) (rather than state anxiety) and administered these measures at the wrong time during the experiment.

This challenge is apparent in the inconsistent findings this research has shown regarding the relationship of language anxiety and task performance in the L2.
Additionally, the effects of learners’ individual factors have usually been studied as a part of the task difficulty component in Robinson’s (2001, 2005) triadic framework. In this sense, both static or trait characteristics of the learner as well as dynamic and variable factors have all been conflated under the same construct, making it extremely difficult to understand the role that each of these individual differences has in isolation.

In what has been considered the dynamic turn, recently, SLA researchers have questioned the traditional way of evaluating language anxiety in the L2 (Dewaele, 2012; Gkonou, Daubney & Dewaele, 2016), highlighted the importance of re-assessing the construct conceptually and empirically. Most of these researchers have considered the ‘complex network of factors impacting on the degree of success of language learning’ (Dewaele, 2012; Gkonou, Daubney & Dewaele, 2016, p. 2). To this dynamic movement in the field of SLA, it is of paramount importance to point out the influence of late proposals, such as the complex dynamics systems theory (CDST) (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008; MacIntyre, Gregersen & Mercer, 2016), which has had a tremendous impact on how affective factors (including anxiety) have been lately understood and re-assessed. Currently, the way certain learners’ characteristics interact with the context and over time has started to attract attention among SLA and TBLT researchers (Baralt & Gurzynski-Weiss, 2011; Gurzynski-Weiss, Jung, Cisneros, DiBartolomeo, Giacomo, Melero-García, Hidalgo, in progress).

Among these context-dependent and unstable variables, language anxiety has been the factor attracting more attention to SLA investigation, yet the one that has posed more scientific and methodological problems (Scovel, 1978).
Another challenge regarding the way language anxiety has been investigated is the little dialog established between SLA (and TBLT) and the psychology and cognitive psychology fields leading to further difficulties in theorizing and measuring this individual characteristic of the learner. In this regard, the psychologist Spielberg (1966) developed the \textit{trait - state} dimensionality of anxiety, the former referring to the individual’s predisposition to experience anxiety in any life situation and the latter associated to the transitory and temporal response triggered by a concrete stimulus.

Additionally, little is known about the different facets of anxiety —the \textit{affective, cognitive, and behavioral}—as defined by Zeidner and Matthews (2011) in an L2 context. Although, a few SLA studies (MacIntyre & Gardner, 1994a; Chen & Chang, 2009) have underscored the detrimental effects of language anxiety on cognitive processes and task performance, this relationship has been underexplored from a TBLT perspective. To this end, understanding what task factors lead to higher levels of language state anxiety (as concurrent to task performance) on learners would certainly shed light on the three-way interface between \textit{Language Anxiety - Cognition – L2 Performance}. To my knowledge, no study has attempted to explicate this interconnectedness by manipulating the cognitive complexity of L2 task.

The three facets of anxiety need to be examined together from a TBLT perspective to see how they unveil and explain the nature of language anxiety in L2 tasks that posed distinct cognitive demands to the learners, and how they subsequently relate to learners’ task performance and linguistic outcomes in an L2 context.

Recent literature on cognitive task complexity (Norris, 2010; Révész, 2014; Sasayama, 2016) has pointed out the need to validate the actual cognitive demands of L2
tasks versus the intended demands. The majority of these researchers have employed cognitive load measures (Sweller, 1994) to investigate the relationship between cognitive load and task complexity in L2 tasks. For validation of the construct of cognitive complexity, to contribute to this strand of research, and to address the connection between state anxiety and cognition, it was decided to incorporate cognitive load measures to the current dissertation.

The present dissertation hopes to illuminate issues connected the TBLT domain that pertain cognitive task complexity and the role of individual differences, more specifically language anxiety, on task performance from the learner’s perspective. The current project aims to provide evidence to better?: a) the nature of language anxiety during oral task performance, as well as which factors contribute the most to induce anxiety during oral production, b) the relationship between cognitive task complexity and state anxiety when cognitive demands are manipulated in the L2, c) the interconnectedness between Language Anxiety—Cognition—L2Performance, d) the relationship between state anxiety and performance as perceived by learners of Spanish, e) the relationship between cognitive task complexity and L2 performance.

By exploring the multiple relationships among affect, cognition and task performance in Spanish as the L2, the current dissertation hopes to aid language instructors and curriculum developers to design second language (L2) tasks that better attend learners’ emotional needs. It is hoped increasing awareness about this three-way interplay will potentially guide language teachers and educators on how to better balance cognitive and emotional needs in L2 tasks. In this regard, it is necessary to design materials that pose the
adequate amount of cognitive demands on L2 learners—both conceptually and linguistically.

The following chapter examines the most relevant theoretical models that have guided the construct of cognitive task complexity and reviews the most influential literature that exists today in this research area. Gaps in this area of research will be discussed, and solutions to address these limitations will be proposed.
CHAPTER 2: COGNITIVE TASK COMPLEXITY

2.1 The Construct of Cognitive Task Complexity

A current debate that has been taking place in the TBLT field for over two decades is related to the process of selecting and sequencing tasks to promote language learning and development (Robinson, 2009). In this regard, SLA researchers (Long, 2015) have pointed out the importance of sequencing and organization of tasks respecting the internal “learner syllabus” (p. 77). This idea led TBLT scholars (Candlin, 1987; Long & Crookes, 1992; Prabhu, 1987) to carefully question and investigate the best way to organize pedagogic tasks within a syllabus taking account the cognitive demands required to perform the tasks. Logically, suggestions were made that tasks should be sequenced following a logical order from the easiest and less complex to the most difficult and complex (Long & Crookes, 1992; Prabhu, 1987). For these early theoretical frameworks, although slight differences were proposed in relation to the most essential aspects that mattered in creating and sequencing pedagogic materials, Prabhu (1987), Candlin (1987) and Long and Crookes (1992) all believed in a progressive increase of cognitive demands for task complexity until reaching the intended target linguistic structures.

Based on these early proposals for task complexity and sequencing decisions, two influential theoretical frameworks have been proposed on task complexity aimed to explain the relationship between feature manipulations in L2 tasks and learners’ cognitive processes (e.g. mainly allocation of attentional resources) when performing a more cognitively complex task, and its subsequent effects on linguistic outcomes and language produced. These frameworks are Robinson’s Cognition Hypothesis (CH) (2001a) and Skehan’s Limited Attentional Capacity Model (LAC) (1998), and they have been
competing hypotheses applied by most of the literature on task complexity to explain their research findings and still remaining unresolved until the present day.

2.2 Robinson’s Cognition Hypothesis and Skehan’s Limited Attentional Capacity Model

Robinson’s Cognition Hypothesis (CH) (2001ab) and Skehan’s Limited Attentional Capacity Model (LAC) (1996, 1998) aim to explain the relationship between task design features and learners’ cognitive processes, mainly, attention allocation during task performance. Importantly, although these models both account for the same phenomenon, they make very distinct predictions on a) how L2 learners interact with cognitive demands, and b) how this relationship between the task and learners’ cognitive activities subsequently affect the different linguistic dimensions of L2 task performance, namely syntactic complexity, accuracy, lexical complexity and fluency (CALF).

Robinson’s (2001a, b) and Skehan’s (1996, 1998) frameworks have been greatly influenced by earlier attentional theories, –Schmidt’s (1990, 1993, 1995) noticing hypothesis and VanPatten (1990) –and first language (L1) models of language acquisition—Givon (1985), and Levelt’s (1989) Speech Production Model, respectively.

Critical theoretical underpinnings of Robinson’s (2001a) CH were derived from Givon’s (1985) L1 model of acquisition. Specifically, Robinson believes that L2 learners follow a similar equivalent pattern of language learning as children do when they acquire their first language. In his acquisition model, Givon (1985) postulated that “linguistic structural complexity tends to accompany functional complexity in discourse and it is therefore reasonable to assume that in many cases increasingly cognitively demanding tasks will make increasingly functional demands in the learner” (cited in Robinson, 2001a:
Robinson (2001a) argues that more complex tasks promote more “pushed and modified output” (p. 206). In his discussion of the role of cognitive task complexity, Robinson’s model is also based on Swain’s (1985) notion of modified output, and Schmidt’s (1990, 1993, 1995) noticing hypothesis and the role of attention. Swain emphasizes how cognitively complex L2 tasks push learners to modify their output by helping them fill the knowledge gap in their interlanguage (Swain & Lapkin, 1995: 372 & 375). Similarly, Schmidt’s (1990) notion of attention is applied to more complex tasks as leading to more noticing of relevant forms in the input, problematic forms in the output, and eventually more incorporation and modification of forms.

One of the most important underpinnings upon which the Cognition Hypothesis is built is Wickens’ (1984, 1992, 2007) notion of attention. In his multiple resource theory, Wickens argues that form and content are not necessarily in competition for attentional pools (Robinson, 2001a), and therefore, no withdrawal of attention occurs when simultaneously attending both. This idea is essential to understanding the CH, which believes that attentional resources are unlimited.

With regards to the prediction of his model on L2 task performance, Robinson’s (2001, 2007) Triadic Componential Framework (TCF) has been proposed as an operational taxonomy to describe task characteristics. The researcher established the existence of three main components in L2 tasks: cognitive task complexity, task difficulty and task conditions. Cognitive task complexity has been defined according to inherent task features, which eventually impact cognitive demands. The second component of the TCF, namely task difficulty, is related to learners’ individual differences and how difficult the task is perceived as judged by learners. Finally, the third component, task conditions, refers to
situation-dependent factors, such as learners’ participation (one-way/two-way tasks), and task participant’s variables (gender, familiarity with the participant, shared knowledge, equality of status, role, among others) (Robinson, 2011).

For *cognitive task complexity*, the TFC has suggested that cognitive demands can be manipulated according to two variables: (a) resource-directing variables and (b) resource-dispersing variables. Robinson’s Cognition Hypothesis assumes that complex tasks along the resource-directing variable dimension will direct learner’s attention to particular aspects of L2 performance, specifically, complexity and accuracy (Robinson, 2007). According to Robinson (2007), resource-directing variables entail conceptual and cognitive demands. Typical resource-directing manipulations involve (a) increasing the number of elements present in the task (i.e., [+/- few elements]), (b) increasing or decreasing special, causal or intentional reasoning demands (i.e., [+/- no reasoning], (c) referring to events closer or farther in space (here versus there) or in time (present versus past) (i.e., [+/- here and now]), or (d) by using a first-person (simpler) versus a second or third person perspective (more complex) (i.e., [+/- perspective taking]). The effects of increasing cognitive task complexity have a direct impact on L2 task performance for the CALF linguistic dimensions by “directing cognitive resources to a wider range of functional and linguistic requirements” (Robinson, 2001a, p. 210). For example, if three characters (versus simply one) are included in a picture narration task, the assumption is that the linguistic forms employed by the learner to refer to those characters will be also more complex.

On the other hand, the resource-dispersing variables involve learners’ performative and procedural demands (Robinson, 2011). Typical ways to manipulate cognitive task
complexity are the following: including the presence of planning time (i.e., [+/- planning time]), altering the task’s prior knowledge (i.e., [+/- prior knowledge]), increasing the number of steps involved for task performance (i.e., [+/- few steps]), presenting the learner with a single versus a dual task (i.e. [+/- single task]). Importantly, the model predicts that making a task more complex along the resource-dispersing demand will lead to depletion of attentional and memory resources to other information processing tasks (Robinson, 2001a). For example, if planning time is provided in a narration task, learners’ attention will be dispersed towards organizing their thoughts in a coherent manner, diverting their cognitive resources from the main task (Robinson, 2011).

As discussed above, the other most widely used theoretical model in cognitive task complexity research has been Skehan’s Limited Attentional Capacity Model (LAC) (1998). Critically, while Robinson’s CH assumes multiple pools of attention resources, Skehan’s follows VanPatten’s (1990) notion of limited single pool of attention. Essentially, the primary assumption of this model is that human resources and working memory capacity are limited (Baddeley, 2007), and therefore, attention to both form and meaning cannot occur at the same time. As opposed to native speakers, non-native speaker learners are forced to use controlled processes in their L2 since most of their knowledge is not automatized. Because of this, when L2 learners perform a complex task, their attentional and memory resources are more focused on the task content and thus, the language produced is affected (Skehan, 2009). Specifically, while the cognition hypothesis does not predict a trade-off effect between accuracy and complexity in learners’ L2 production; Skehan’s model (1996), due to human's limited attentional resources, predicts such trade-
off to occur, making it difficult to simultaneously attend to accuracy and complexity during task performance.

Using a different approach than the CH, Skehan (1996) explains cognitive task complexity according to three main components: (a) cognitive complexity, (b) code complexity and (c) communicative stress. According to LAC, cognitive complexity is determined by two principle aspects, namely processing and familiarity. In short, processing is related to the intrinsic cognitive demands (Sweller, 1994) imposed by the content of the task during task performance. However, familiarity has to do with how much knowledge or information learners actually possess about the topic of the task. Code complexity is associated with the level of difficulty of the language forms required by a task for successful performance. Finally, communicative stress has been described as the pressure and situational anxiety that occurs as a result to communicate in the L2. In terms of cognitive complexity, the higher the processing, the higher the cognitive complexity, whereas the more familiarity with the content of the task, the simpler the task becomes (Skehan, 1996). The model predicts that in more complex tasks (for instance, when increasing the number of characters present) online processing will lead to a reduction of attentional resources available to attend to language production. Therefore, a trade-off between linguistic complexity and accuracy in the L2 takes place since learners cannot attend to both dimensions simultaneously due to the limited attentional resources.

In addition, Skehan has also relied on Levelt’s (1989) Speech Production Model to explain learners’ allocation of attentional resources during task performance. In this first language acquisition model, three production stages take place in speech. First, in the conceptualization stage, idea organization and development of the message takes place,
followed by the formulation stage, where ideas get encoded into lexical and syntactic structures, and finally, the formulation stage consists of the phonological realization of the output (Skehan, Xiaoyue, Qian, & Wang, 2012). Importantly, Skehan (2009) and Skehan et al. (2012) have made connections between L1 and L2 output in how task design and task conditions can be manipulated to enhance speech production in L2 task performance.

Although both the CH and the LAC models have been widely applied by previous literature in the TBLT area, they fundamentally disagree on the effects that cognitive complexity has on L2 performance, development and learning. Whether attention can be paid simultaneously to different aspects of the language, namely complexity and accuracy, is still a matter of debate today.

Another of the most striking aspects about previous literature devoted to the role of task complexity is the many distinct and various operationalizations of the construct that exist to date. A synthesis study (Sasayama, Malicka & Norris, forthcoming) revealed the existence of over 129 studies published in refereed journals. Specifically, according to Sasayama (2016), there are at least 14 different ways in which this construct has been operationalized based on Robinson’s CH and Skehan’s LAC. In sum, despite the many studies that have investigated the effects of cognitive task complexity on L2 linguistic outcomes so far, findings have been mixed and inconsistent. The following section reviews the most important strands of research within the task complexity literature and their findings on L2 performance measures.

2.3 Cognitive Task Complexity Research

Cognitive task complexity has recently become an area of interest in the area of task-based language learning and SLA research. A vast number of previous studies (Albert,
2011; Baralt, 2013; Dörnyei & Kormos, 2000; Gilabert, 2007a, 2007b, 2009; Ishikawa, 2006; Kim & Tracy-Ventura, 2011; Kuiken & Vedder, 2008; Kormos & Trebits, 2011; Lee, 2002; Révész, 2011; Révész, Sachs & Hama, 2014; Révész, Michel, Gilabert, 2016; Robinson, 2001, 2007, 2011; Sasayama, 2016; Sasayama & Izumi, 2012; Torres, 2013, Yoshida, 2012) have examined its effects on L2 performance, L2 development and L2 learning. Importantly, one of the reasons why cognitive task complexity has been widely investigated is that it allows researchers to operationalize cognitive demands in tasks (Torres, 2013).

Following Sasayama et al.’s (forthcoming) synthesis study, Table 1 shows the most common operationalizations for cognitive task complexity¹ in order of relevance from the most to the least commonly used:

**Table 1:** Most common operationalizations of task complexity.

<table>
<thead>
<tr>
<th>Operationalization of task complexity</th>
<th>Studies</th>
</tr>
</thead>
</table>
Table 1. (Cont’d)

<table>
<thead>
<tr>
<th>Operationalization of task complexity</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>+/-Simple storyline</td>
<td>Tavakoli &amp; Foster (2008), Tavakoli (2009)</td>
</tr>
<tr>
<td>+/-Time pressure</td>
<td>Skehan &amp; Foster (1999), Maad (2010)</td>
</tr>
<tr>
<td>+/-Perspective taking</td>
<td>Iwashita et al. (2001), Elder et al. (2002)</td>
</tr>
<tr>
<td>+/-Integrated format</td>
<td>Al-Shehri &amp; Gitsaki (2010)</td>
</tr>
</tbody>
</table>

1 Note: *These studies combined manipulation of +/- few elements with +/- reasoning.*
(Summary of task complexity studies, adapted from Sasayama, 2015)
First, as can be seen in Table 1, it is important to note that in all previous studies task complexity was operationalized with two levels (simple vs complex) of higher and lower cognitive demands. Second, the different manipulations of task features have yielded distinct effects of each task design on cognitive complexity and task performance for L2 learners (Sasayama, 2015), and therefore, comparisons among studies on each operationalization cannot be made due to the very different operationalizations. For the purpose of the current study, a review of the literature on the two most common operationalizations, +/- planning and +/- few elements will be presented for comparability reasons.

The most prominent areas of research of cognitive task complexity that have led to the most consistent results have examined how the provision of planning time (+/- planning) has affected learners’ L2 task production or performance. The majority of studies in this strand of research used a narrative task and operationalized task complexity giving either pre-task planning (provided prior to task performance) or online planning (during task performance). For pre-task planning studies, the most popular length of time for planning provided has been 10 minutes, and overall, findings have been very consistent for the effects of planning on CALF measures. In particular, previous literature on the effects of pre-task planning on oral production (Foster & Skehan, 1996; Ortega, 1999; Iwashita, McNamara & Elder, 2001; Iwashita & McNamara, 2002; Mochizuki & Ortega, 2008; Tavares, 2009; Yuan & Ellis, 2003) and written performance (Ellis & Yuan, 2004; Farajani & Meraji, 2011; Meraji, 2011; Mohammadzadeh, Dabaghi & Tavaloki, 2013) yielded that pre-task planning promotes more fluent and more complex language (Ellis & Yuan, 2004; Farajani & Meraji, 2011, Foster & Skehan, 1996; Gilabert, 2007a; Meraji, 2011; Ortega,
1999; Sasayama & Izumi, 2012; Yuan & Ellis, 2003). For accuracy, whereas some studies on speech production, results revealed either a positive effect (Foster & Skehan, 1996; Ortega, 1999; Mochizuki & Ortega, 2008; Tavares, 2009) others showed a lack of effect (Foster & Skehan, 1996; Yuan & Ellis, 2003). Interestingly, those studies that examined lexical variety both on oral and written production found no statistical significant difference between task complexity levels for this measure (Ellis & Yuan, 2004; Farajani & Meraji, 2011; Meraji, 2011; Yuan & Ellis, 2003). The following section reviews previous literature on cognitive task complexity that has used the +/- few elements as the primary feature of cognitive demands manipulation in task design.

### 2.4 Previous Research Operationalizing Task Complexity Via the Number of Elements on the Resource-Directing Dimension in Robinson’s Cognition Hypothesis

Table 1 above shows that the second most popular operationalization of cognitive task complexity in TBLT has been based on Robinson’s resource-directing dimension of cognitive resources, namely, +/- few elements. Sometimes, these were manipulated alone, and often times they were combined with reasoning demands manipulation. Below, the most relevant literature that employed number of elements following Robinson’s operationalization’ of cognitive task complexity is reviewed. Findings related to individual differences or level of proficiency pertaining the following studies are only briefly discussed or omitted, since they are reviewed separately and more in depth in the subsequent individual differences and L2 proficiency sections.

An early study, Robinson (2001b) investigated cognitive task complexity manipulating +/- few elements in a dialogic map-task with two levels of syntactic complexity (simple vs complex task). 22 English learners, whose L1 was Japanese were
assigned to a task role, namely, information giver or receiver. Specifically, information giver had to provide the interlocutor with directions to get to a destination point, and the information receiver role was allowed to ask clarification questions. The study examined language performance on CALF measures for the information-givers and amount of negotiation by the information receivers. Interestingly, results showed the complex task promoted lexical variety for the information givers and a higher number of clarification and confirmation requests for the information receivers. However, a neutral effect was found for task complexity on the complexity, accuracy, and fluency (CAF) measures.

In Lee (2002), 82 learners of Korean participated in two dialogic picture description tasks where +/- reasoning and number of elements was used to make the task more cognitively complex. There were three levels of task complexity (simple, complex and more complex), and the story-telling tasks were about a car accident and the human body. For the accident task, number of cars, direction of the cars, pedestrians and road conditions were manipulated. For the human body task, the different versions of the task differed in a) stages of the body system manipulated—blood circulation, water intake and loss, and movement in the intestine, b) arrows of movement, and c) the number of words provided in a vocabulary list of terms translated from Korean to English. No positive effect was found for accuracy or syntactic complexity, and the researcher finally concluded that results did not support the CH. Lee recommended the use of reasoning demands as main manipulation feature of task complexity and provided pedagogical implications for it in future task design.

Kuiken and Vedder (2007) examined task complexity in L2 writing performance. Reasoning and +/- few elements were combined as main features of cognitive complexity
manipulation. For Kuiken and Vedder (2007) participants were 91 Dutch learners of Italian and 76 Dutch learners of French that differed in their L2 proficiency level, low to advanced level of proficiency. Participants had to write a letter recommending a travel destination to a friend, in either French or Italian, that varied in the number of requirements (3 versus 6). While, the complex task yielded an increase in accuracy and lexical variety in the written output as compared to the simple task, no effect was observed for lexical syntactic complexity between the simple and the complex version of the tasks. Kuiken and Vedder (2008) replicated their study using the same research design and looking at the same outcome measures. This time participants were 84 L2 Italian and 75 French learners, and researchers employed more specific measures for accuracy and lexical variation to look at what specific type of errors learners produced and distinguish high frequency from low frequency words. Results indicated that for the complex task the improvement on accuracy was shown particularly for lexical errors. For lexical variation, mixed results were found where learners of Italian produced a higher number of frequent words in the complex task, and French learners a higher number of infrequent words in the complex task. Although the researchers concluded that the study supported Robinson’s CH, no relationship was revealed between L2 proficiency and task complexity. The researchers conducted an extension of the previous studies, Kuiken & Vedder (2011), where they examined task modality—oral versus written—on L2 performance measures using the same previous task manipulated by +/- few elements. This time participants were only 44 L2 Italian adults assigned to the oral mode that were compared to the 91 participants from the researchers’ previous study on the written mode. Based on a cloze test, participants were divided into low and high proficiency groups. Results revealed task complexity led to higher levels of
accuracy on both oral and written task modes. Interestingly, for the written mode no differences between tasks were revealed for syntactic complexity. However, and contrary to the CH, for the oral mode a negative effect was found for the more cognitively complex task on syntactic complexity.

In another study, Michel, Kuiken and Vedder (2007), cognitive task complexity was manipulated via +/- number of elements and task conditions (monologic versus dialogic tasks). Forty-four learners of Dutch were randomly assigned to either the monologic or dialogic task condition group, and all participants performed the two tasks—the simple and the complex versions of the task—under a within-group design. All participants were provided with a leaflet where MP3 players or mobile phones were for sale. For the monologic task, learners had to leave a voice mail to a friend giving advice on what product to buy, and for the dialogic version, they had a conversation over the phone with a friend to decide which device they would buy. To examine the effects of task complexity on task performance, CAF measures were analyzed. Results showed that a positive effect for accuracy in terms of total number of errors and repairs for the complex task. Therefore, an increase of cognitive demands seemed to produce more accurate language. However, no differences were found for fluency or syntactic complexity between the two tasks. Although the complex task in the monologic condition led to higher accuracy ratio as compared to the dialogic task condition, no significant interaction was found between task complexity and task conditions.

Gilabert (2007b) investigated the impact of cognitive task complexity on self-repairs, an accuracy measure, on 42 learners of English. Participants were enrolled in two different Catalan institutions and were assigned to a low proficiency (N=29) and high
proficiency (N=13) group based on a placement test. They were randomly assigned to one out of three different task types (a narrative, an instruction-giving, and a decision-making task) and completed the simple and the complex task. Task complexity was manipulated for the narrative task along +/- Here-and-Now, for the instruction-giving task along +/- few elements and for the decision-making task along +/- reasoning. As the researcher predicted, the number of self-repairs increased for the complex task. No differences were found between the low and the high proficiency levels in relation to the number of self-repairs produced.

Gilabert, Barón and Llanes (2009) hypothesized an increase in cognitive task complexity would lead to more opportunities for interaction for learners. Based on Mackey and Abbuhl (2005), interaction was measured by comparing language related episode LREs and recasts (less explicit type of feedback) to see if more negotiation of meaning took place in the complex task. 6 different tasks were employed and task type –narrative, instruction-giving decision-making—was manipulated so that participants received the simple and complex version of one of the three task types. Participants were 60 learners of English as an L2. Task complexity was manipulated differently for each type of task. For the narrative task, +/- here and now was used, for the instruction-giving task number of elements differed between the simple and the complex task, and finally, for the decision-making task, participants had to engage in different amounts of reasoning. To examine task performance, self-repairs were analyzed to see if these were LREs, recasts or examples of negotiation of meaning. Findings confirmed the researchers’ hypothesis showing that the complex task indeed led to more interaction opportunities as measured by the total number of negotiation for meaning episodes. Additionally, task type played a significant role since
the advantage for language interaction was observed for the narrative and the instruction-giving tasks.

Kim (2009) also explored the effects of task complexity on oral production of LREs, considering task type and level of proficiency as moderating variables. 34 English learners were assigned to high and low proficiency levels based on institutional course enrollment and the TOEFL exam. The study employed a picture narration task manipulated by reasoning demands, and a picture difference task manipulated by +/- few elements. All participants performed 4 tasks, the simple and complex version of the narrative task, and the simple and complex version of the picture difference task. Interestingly, results showed that the type of feature manipulated for task complexity had different effects on learners’ performances according to their level of proficiency. For the high proficiency learners, the complex with higher increasing demands task yielded a statistical significant increase for LREs. Importantly, however, the same task for low proficient learners revealed the opposite results showing a significant lower number of LREs for the complex task when compared to the simple task. Critically, when number of elements was manipulated for the picture difference task, beneficial effects for the cognitively complex task were observed for the low proficient learners this time, and only a small advantage was found for the high proficiency groups in this case. This study had important pedagogical implications in that manipulation of different task features for cognitive task complexity affected learners differently depending on their proficiency level.

Révész (2011), in a classroom-based study, also examined the effects of task complexity on interaction—LREs—and on L2 accuracy and complexity. Like all the previous studies, she used +/- few elements and reasoning demands combined together on
an argumentative task. 43 English learners from an intermediate and advanced level of proficiency were randomly assigned to the simple or the complex task and their task performance compared. Results revealed that for the complex task learners’ output was more lexically varied and accurate, and contained a higher number of LREs. However, no advantage was found for syntactic complexity in the complex task.

Michel (2011) investigated how task complexity affected interaction and linguistic outcome measures, the CAF triad. 64 Dutch learners were assigned to either a monologic or dialogic condition and performed a decision-making task. The task consisted on deciding on a female/male dating couple and a Belgium-Dutch studying peer. In addition, tasks were manipulated in +/- few elements, the simple task containing 4 candidates and 4 characteristics, and the complex version containing 6 candidates and 6 characteristics. All participants completed the simple and complex version in each condition. Results revealed that the complex task promoted lexical diversity in monologic and dialogic task conditions. No effects were found in the other conditions.

Kim and Tracy-Ventura (2011) focused on whether task complexity via reasoning demands had an effect on L2 development of past tense morphology. Participants were 82 university learners of English (from a beginning to an upper intermediate level of proficiency). The research design included three levels of complexity (simple, complex and very complex) manipulated using both reasoning demands and +/- few elements. Participants performed a communicative task in pairs where they had to talk about their college experiences, university orientation, and campaign preparation. To examine L2 development of past tense, two assessment measures were employed using a pre-, post-test and delayed post-test design. One of the measures was an individual picture description
oral task in the past tense, and the other was a pair oral production task of true/false activities. Findings of the study showed that the complex and the very complex groups outperformed the simple group in all assessment measures in the post-test and delayed post-test.

In another interaction study, Kim (2012) addressed how task complexity affected the amount of LREs opportunities and L2 development of English question formation. Participants were 191 L1 Korean learners of English with varying proficiency levels from beginning to intermediate (based on the TOEIC exam). The experimental conditions were manipulated by increasing +/- elements and +/-reasoning, and participants were randomly assigned to one a simple, a + complex or a ++ complex task. L2 performance was measured through comparing learning opportunities (LREs), and question development was examined by using one individual and two pair oral assessment tasks in a pre-post-test, delayed post-test design. Results showed that the ++ condition produced a statistically significantly higher number of LREs and an increase in the stages of question development.

Sasayama (2011) tested Robinson’s CH by looking at the impact of cognitive task complexity on learners’ oral and written performance. Participants were 10 English learners whose proficiency scores on their iBT TOEFL test ranged from 68 to 98 (out of 120). The study employed four monologic narrative tasks, two written (simple vs complex) and two oral (simple vs complex), and all participants performed the four tasks. Task complexity was manipulated via +/- few elements in cartoon-based stories. The complex task had a higher number of characters, foreground and background elements, while the simple task did not have background elements and presented fewer characters and foreground elements. Results revealed the complex task in the written mode showed more
complex language as compared to the speaking complex version. Additionally, a neutral effect on accuracy was found. In interpreting the results, the researcher concluded a) the cognition hypothesis seemed to be more relevant for language complexity than accuracy, and b) accuracy seemed to be more dependent on individual abilities of learners on producing accurate constructions than on the inherent cognitive demands of tasks (Sasayama, 2011). In light of the findings, it was concluded that the study partially supported Robinson’s CH, and partially disconfirmed it.

Sasayama and Izumi (2012) investigated the effects of task complexity on oral production on Japanese learners of English. Cognitive demands were manipulated along both the resource-directing and resource-dispersing line. In particular, the study operationalized task complexity via +/- few elements and the provision of pre-task planning. Twenty- three English high school learners received two narrative tasks, which differed again in the number of characters present in the cartoon-based pictures. Out of the 23 participants, 10 received pre-task planning and the other 13 did not receive any planning. Findings revealed that the more complex task led to higher complexity levels. However, this time global accuracy and fluency were negatively affected. Additionally, pre-task planning benefited learners in relation to syntactic complexity, but had a negative effect on fluency. Critically, the researchers concluded that results partially supported and partially disconfirmed both Robinson’s CH and Skehan’s LAC.

2.5 Current Limitations in the Operationalization of Cognitive Task Complexity

Findings from the majority of the studies reviewed above (with the exception of the +/- planning literature) have shown unclear and mixed findings on the effects of cognitive task complexity on the different CALF L2 performance dimensions. As a consequence, a
few recent studies (Révész et al. 2016; Sasayama, 2016) have emphasized that the many different operationalizations of cognitive task complexity across the existing TBLT literature today have made it impossible to make predictions regarding its effects on L2 performance.

Specifically, recent concerns have been recently raised about the general assumptions on the cognitive demands imposed by L2 tasks in previous SLA literature (Sasayama, 2016). In other words, there has been a tendency among task complexity researchers of focusing on the linguistic outcomes observed, to later draw conclusions about the cognitive demands that make the tasks more or less complex. In this respect, although, overall, more complex tasks have been shown to promote L2 performance and development (Robinson, 2011), scant SLA research to date has demonstrated this empirically, considering “what” in particular makes a task more or less cognitively complex. To date, only a few recent TBLT studies (Révész, Michel & Gilabert, 2016; Sasayama, 2016) have aimed to validate the construct of cognitive task complexity by employing various cognitive load methodologies.

Additionally, based on cognitive psychology research (Brünken, Plass, and Leutner, 2003; Paas & van Merriënboer, 1994; van Gog & Paas, 2008), SLA and TBLT researchers (Norris, 2010; Norris & Ortega, 2003; Révész, 2014; Révész, Michel & Gilabert, 2016; Sasayama, 2016) have lately claimed the need to examine the cognitive load that pedagogic tasks impose on language learners by means of instruments that verify and measure the actual cognitive demands of L2 tasks.
2.6 Cognitive Load Theory and Mental Effort in Educational Research

Sweller’s Cognitive Load Theory (CLT) (1994) from cognitive psychology has been widely used in educational research and takes into careful consideration learners’ limited working memory capacity (Baddeley, 2007). One of the primary goals of the application of this theory was designing pedagogic materials that would promote learning in instructional settings by considering excessive burden of learners’ working memory capacity.

CLT established three sources of cognitive load: intrinsic, extraneous, and germane (Sweller, 1994). Intrinsic cognitive load refers to the cognitive demands that result from content of the pedagogic material, which is associated with the learner’s previous knowledge of that content. Extraneous cognitive load is related to the undesired type of cognitive load that takes up cognitive resources but, eventually, does not lead to learning outcomes. An example of this would be learners attending to an irrelevant and unrelated task to the content material to be learned during instruction, delaying actual learning. Finally, germane or facilitative cognitive load refers to learners’ working memory resources that are available and activated when learners are presented with the materials (Sweller, 1994). In short, while the first two types of cognitive load deal with the task itself (intrinsic) and how it is presented (extraneous), the latter (germane or facilitative) is essentially related to the learner’s individual factors in their working memory.

A fundamental distinction needs to be made between the terms cognitive load and mental effort. Paas and van Merriënboer (1994) differentiated between the two concepts. Specifically, while the former has been referred to as the cognitive inherent task demands (cognitive complexity) – invariable elements within the task (Skehan, 1996) – the latter is
defined by a conscious and intentional effort exerted by learners during task performance (P. 420). Additionally, the researchers highlighted that mental effort was dependent on the following causal factors: task environmental characteristics, subject’ characteristics, and interactions between these two.

Specifically, Paas and van Merriënboeret (1994) defined mental effort as “the amount of resources actually allocated to accommodate the task demands” and “the amount of controlled processes in which the individual is engaged” (p. 420). Importantly while cognitive load is attributed exclusively to the task, mental effort is more related to task difficulty in that both are dependent on the learner’s characteristics and perception. It could be said that mental effort refers to how a learner reacts to the cognitive demands imposed by the task, and how much he/she is willing to process the input more or less deeply, which will consequently affect their subsequent performance. For example, in a very complex task a learner can decide willingly to barely “cognitively” engage in the task (despite he/she could do otherwise) because he or she may be discouraged and perceived it as too difficult (task difficulty). In this case, the learner could give up and dedicate almost no cognitive resources, probably resulting in a poorer performance had he/she put more mental effort in the task.

In sum, although cognitive load and mental effort are related to each other, differentiating the two is essential to understand how each of them contribute to explain cognitive task complexity.
2.7 Task-Based Language Teaching Research and Independent Cognitive Complexity Measures

The striking number of mixed findings in cognitive task complexity literature that exist today reflects important shortcomings in how the construct has been operationalized, leading to an increasing interest in the SLA research to look outside the field to find answers (Norris & Ortega, 2003; Norris, 2003; Sasayama, 2016). These researchers have acknowledged the importance to incorporate various distinct independent cognitive load methodologies (Brünken et al., 2003) to validate the “supposed” cognitive demands intended for designed-to-be simple versus designed-to-be complex tasks.

This section briefly reviews some of the most widely employed cognitive load measures and their contributions to tap cognitive task complexity on L2 tasks. In particular, in TBLT research the most important measures used have been task difficulty (Robinson, 2001b; Gilabert, 2007; Révész et al., 2016; Shiau & Adams, 2011; Sasayama, 2016), mental effort (Révész et al., 2016; Sasayama, 2016), eye-tracking (Révész, Sachs & Sama, 2014), dual-task methodology (Révész et al., 2014, Révész, Michel & Gilabert, 2016, Sasayama, 2016), time estimation (Baralt, 2013), expert judgment (Brown et al., 2002; Révész et al., 2014; Révész et al., 2016), and introspective methods, such as stimulated recall and interviews (Kim et al., 2015; Tavaloki, 2009). For the purpose of the current study, the literature pertaining to task difficulty, mental effort, and time estimation will be discussed as these are the measures selected for the current study.

2.7.1 Task Difficulty and Mental Effort Measures

Self-report of difficulty or task difficulty self-assessment – categorized as subjective direct type measures (Brünken et al., 2003) – have been by far the most widely
used in the TBLT domain. Following Kalyuga, Chandler and Sweller (1999), self-report measures of difficulty require participants to estimate on a Likert scale—from 1 (extremely easiest) to 9 (extremely difficult)—how difficult they perceive the task to be. As a subjective measure, it is important to note task difficulty self-assessment does not provide a direct indication of cognitive load, but rather taps the cognitive processes involved in a task as perceived and judged by participants. In this sense, mediating factors other than the task alone (learners’ language ability and competency) (Van Gog & Paas, 2008) could likely interfere in the ratings provided by learners.

As opposed to task difficulty, Brünken et al., (2003) categorized mental effort as a subjective indirect measure since they argue that the causal relationship between the results of the measure and the actual cognitive load of the task is not direct. The researchers pointed out that whereas perceived difficulty is a direct result of the cognitive load imposed by a particular task, mental effort may not be necessarily (although it could be) a consequence of the cognitive load, but instead a strategy employed by learners to mentally cope with higher cognitive load tasks.

### 2.7.2 Introspective Methods

A variety of introspective methods in SLA and TBLT research has been employed to tap on learners’ internal processes and inform researchers of how learners’ divide their attention and process task input. One of the most popular and prominent introspective methods employed today to examine cognitive processes is stimulated recall. For stimulated recalls, learners are asked to report what they were thinking at the time of their performance based on either video or audio stimuli (Egi, 2010; Gass & Mackey, 2000). Importantly, a stimulated recall session must be performed immediately after performance
to avoid learners’ forgetting about their thoughts during task performance. Interviews are another introspective method, and SLA researchers use them as they reveal “participant’s meanings and interpretations of their experiences and actions” (Dörnyei, 2007: 38).

2.7.3 Dual-Task Methodology

Dual-task methodology has been considered a direct objective methodology of cognitive load and it has been used in cognitive psychology (Brünken et al., 2003). This measure involves performing a secondary task parallel to a primary task. It has been considered a direct objective measure because it measures cognitive load concurrently to task performance as opposed to self-rating measures that are subjective measures (Gilabert et al., 2016). While the primary task is the main pedagogical task of interest for the study, the secondary task consists of a repetitive simple task that requires the learner’s continuous attention (Brünken, Steinbacher, Plass, and Leutner, 2002). Usually this task will consist of simply noticing an auditory (Brünken, Plass, & Leutner, 2004) or a visual (Cierniak, Scheiter, & Gerjets, 2009) stimulus. The assumption of this methodology is that the amount of cognitive load of the primary task will be reflected in how fast learners react to the secondary task (Brünken et al., 2003). Performance in the secondary task is usually measured by reaction times and the learner’s accuracy on the task (Révézs, et al, 2016). A task that poses higher cognitive load will generally result in longer reaction times and a less accurate speech.

2.7.4 Time Estimation

Another measure used in cognitive psychology to account for cognitive load has been time estimation (Block, Hancock and Zakay, 2010; Brünken et al., 2003). The premise for this measure is that the perception of time varies depending to the degree of cognitive
demands imposed by the task on the learner. As argued by Woehrle and Magliano (2012), “allocation of attention affects the awareness of time passage, particularly during a cognitive task” (p. 314). Two approaches have been employed for this method: the retrospective and the prospective approach. For the prospective approach, when participants are asked to estimate prospectively (prior to task performance) how long they think the task will take, the judged time tends to decrease when the task is more cognitively complex (Block et al, 2010). However, when the task poses low cognitive load, participants usually estimate the time will be shorter. Similar to the dual-task methodology, under the prospective approach when a task imposes high cognitive load, less attentional resources are available to attend to a secondary task, in this case, time judgment. To account for this phenomenon, Zakay and Block (1997) proposed the attentional gate model (AGT), holds that “a person may divide attentional resources between attending to external events and attending to time. Attending to time opens the attentional gate thereby allowing pulses to pass through stream to the cognitive center” (p. 14).

On the contrary, under the retrospective approach, participants are asked to judge the time they think they spent during a specific task, and it performed retrospectively after task completion (Block et al, 2010). This approach is accounted by the contextual change hypothesis (Block, 1978), which claims that retrospective judgments of time depend on retrieving information stored in memory about contextual changes that took place during task performance. These contextual changes refer to changes in environment context, mood, and mental processing that were happening during performance. Zakay and Block, (1997) argue that the rationale behind this hypothesis is that when a task is more cognitively complex, those contextual changes in processing are remembered by learners as actually
taking significantly longer than they actually did. The perception of time is positively correlated with the amount of contextual changes experienced and stored in memory. For the retrospective approach, when the cognitive load of a given task increases, the length of time perceived to have passed also increases.

2.8 Task-Based Language Teaching Research on Cognitive Load Measures

Self-assessed task difficulty has been one of the most popular cognitive task complexity measure employed until today by TBLT researchers (Baralt, 2013; Gilabert, 2006, 2007; Gilabert & Barón, 2013; Gilabert et al., 2009; Gilabert et al., 2011; Kim, 2009; Kim & Tracy-Ventura, 2011; Levkina & Gilabert, 2012, 2014; Malicka & Levkina, 2012; Michel, 2011; Révész, 2009, 2011; Révész et al., 2016; Shiau & Adams, 2011; Robinson, 2001b; Sasayama, 2016). It has also been regarded as one of the most reliable and valid measures used by SLA researchers (Sasayama, 2016). All the studies cited above have adopted Robinson’s (2001b) original task difficulty questionnaire in their studies to evaluate this construct. In addition to task difficulty, this scale included self-ratings for perceived stress, motivation, interest in the task, and ability to successfully perform the task.

To date, two studies in TBLT research (Sasayama, 2016; Révész et al., 2016) have investigated cognitive load indirectly by means of mental effort self-report measures. In addition, and as discussed above, these researchers have also attempted to “validate” the construct of task complexity by combining several of the aforementioned methodologies.

Sasayama (2016) based on her 2015 dissertation was the first to attempt to validate task complexity by providing a conceptualization and operationalization of the construct based on a cognitive psychology theory. The researcher applied Sweller’s Cognitive Load
Theory (CLT) (1994) theoretical underpinnings to task design to investigate whether L2 tasks indeed imposed the complex cognitive demands intended by researchers, proving such complexity empirically (Sasayama, 2016). Sasayama designed her tasks by applying Sweller’s (1994) distinction between the three types of cognitive load present in tasks – extraneous, germane and intrinsic – and discussed how these could be incorporated into task design and promote L2 learning.

Specifically, the study investigated a) whether differences among various task complexity levels are detectable by the aforementioned cognitive load measures (dual-task methodology, time estimation, and self-ratings), b) what exactly makes a task more or less complex, c) how actual cognitive task complexity—as measures by cognitive load measures—is related to L2 performance across the different L2 proficiency levels. In relation to L2 performance, the study also examined what particular measures for fluency, accuracy, syntactic and lexical complexity were better predictors of learners’ performance for the different task complexity conditions. Lastly, with regard to L2 proficiency, the study explored the moderating effect of L2 proficiency on the actual cognitive task complexity of L2 tasks and L2 performance.

One of the study’s most important contributions was precisely the employment of a variety of measures from cognitive psychology to tap cognitive load and validate the assumed cognitive demands imposed by L2 tasks. Specifically, the measures employed were dual-task methodology, retrospective time estimation, and self-ratings for perceived difficulty and mental effort. Participants, 61 native speakers (baseline) and 120 Japanese English learners from different proficiency levels, completed four picture description tasks. Tasks were manipulated along +/- few elements. Specifically, number of characters was
used as the primary feature to manipulate cognitive task complexity, Task 1 (the simplest) containing only 1 character and Task 4 (the most complex) containing 9 characters. Results of the study revealed that the number of elements was one of the main components indeed affecting cognitive task complexity. Overall, cognitive load measures showed that for the native speakers the four tasks posed very similar cognitive demands. For the L2 learners, the four tasks posed distinct cognitive load demands. Importantly though, only statistically significant differences were found between the simplest task (only one character) and the most complex task (9 characters), where only large differences in cognitive demands were more detectable. These results demonstrated that these cognitive load measures were more sensitive when differences in the number of elements between tasks were greater. This supported Robinson’s (2001a) prediction for the positive relationship between increasing number of elements and cognitive demands.

Additionally, findings showed that the type of cognitive load (whether facilitative, intrinsic or extraneous) made a difference in the cognitive demands imposed by tasks. Specifically, results revealed that conceptual input, code complexity and performance factors led to both facilitative and extraneous cognitive load. In this respect, Sasayama (2016) highlighted that to promote L2 performance and learning it was essential to increase the facilitative type of cognitive load and reduce the detrimental task-irrelevant type of cognitive load in L2 tasks (called extraneous cognitive load). To this end, the researcher recommended paying attention to the balance among the three types of cognitive load in task design to maximize beneficial effects of task complexity on L2 linguistic performance. Additionally, the study strongly recommended that triangulation of several cognitive complexity measures was essential to be able to draw conclusions on how cognitive task
complexity affects learners’ language performance. Lastly, results for the effects of L2 proficiency showed that high proficiency learners outperformed the low proficient learners in the most cognitively complex task in most linguistic CALF measures.

In this respect, one of the major contributions of the study for TBLT was the validation of the construct of task complexity by using several independent methodologies in the same research design to determine their reliability for future SLA research. The study concluded that only by determining the learner’s cognitive load during task performance, could conclusions be actually drawn about the presumed cognitive demands imposed on L2 learners (Sasayama, 2016). The researcher highlighted that a combination of different measures of cognitive load or cognitive complexity (via data triangulation) would lead to more robust and solid designs in the operationalization of cognitive task complexity, since interpreting the same data from different angles could only be achieved by using more than one different measures (Kim et al., 2015, Sasayama, 2016).

Révész et al. (2014) investigated manipulation of cognitive task demands (task complexity) on the acquisition of L2 past counterfactual constructions when recasts are provided in communicative meaning-based activities. To assess the validity of cognitive task complexity, the study employed dual task methodology and expert judgment. Participants were 51 L2 learners of English and the study used a pre-test post-test design. For the treatment task, participants first read two texts about the biography of two famous figures, and after reading, considered possible causes for these figures' biographical events. The simple and the complex task were manipulated along the reasoning demands, making the potential cause of the events more apparent in the simple task, and vaguer in the complex version of the task. For the reasoning stage, participants were asked to formulate
counter-factual events of the following kind: “if not, then…but” (e.g. if Eleanor had not learned about the refugees’ problems, then she had not helped to pass laws about the human rights). For the expert judgment measure, researchers asked two applied linguistics to rate the difficulty of 32 counter-factual instances on a 5 point-Likert scale from both the simple and the complex task. An oral production test and two written receptive tests were used to assess linguistic outcomes for the targeted items. For the dual-task methodology, while the primary task was the sentence production of past counter-factual constructions, the secondary task was a color change of the background of the computer screen for a duration of 250 ms. randomly displayed at intervals of 2,500 ms. Both accuracy and reaction times were recorded by E-Prime 2.0. Results for dual-task methodology revealed that participants in the simple task conditions were more accurate as compared with the complex task, suggesting that the cognitive load for the latter was indeed greater. However, contrary to expectation, no differences between the two levels of complexity were found for reactions times for the secondary task. For expert judgment, the two applied linguists rated all targeted items for the task ‘designed as complex’ to be much more complex than the items task ‘designed as simple’. Results showed higher oral production gains for the cognitively simple task over the complex task.

Another recent study, Révész et al. (2016) combined dual task methodology, subjective self-ratings and expert judgments to assess cognitive task demands to validate the manipulated task features. Participants were 96 English learners and 61 English as a second language (ESL) teachers. For the English learners’ group, 48 were native speakers of English, and the other 48 were L2 English learners. All participants performed a simple and a complex version of three tasks: a picture description, a decision-making and a map
task. Additionally, 24 native speakers and 24 L2 learners (half of all participants) took part in the dual-task condition, while the other half did it in the single task condition. For the dual task methodology, the narrative and the decision-making tasks were employed as primary tasks, and the secondary task was again the screen color background change since it had proved to be a sensitive task in Révész et al. (2014) study. For the self-assessment measures, participants were asked to rate their perceived mental effort and task difficulty, and for the expert judgment, teachers had to evaluate the complexity of the tasks. Results for dual-task methodology supported again the same patterns reported in Révész et al. (2014) showing significantly lower levels of accuracy for the cognitively complex task, but no statistical differences for reaction times between the simple and the complex task. With the exception of the narrative task, for the other two task types, the cognitively complex task was judged by learners as being statistically significantly more complex when compared to the simple task, confirming the researchers’ predictions. Lastly, for the expert judgment measure, statistically significant differences were both revealed in the ratings of mental effort and task difficulty for teachers’ ratings, also validating the intended levels of cognitive task complexity.

For introspective methodologies, Tavaloki (2009) investigated task difficulty by performing retrospective interviews on 10 intermediate English learners and ESL instructors teaching at the same college as the learners selected for the study. Four oral narrative tasks from Tavaloki and Skehan (2005) were used and manipulated by the structure of the narrative and the timeline. All participants performed the four narrative tasks, and after completion were interviewed to get insight on fact factors and the reasons why certain tasks were perceived as difficult for them. After examining the same tasks, the
language instructors were interviewed. Results revealed although some differences existed between learners’ and teachers’ perspectives, overall the two groups mostly agreed on what factors made the tasks more difficult. Among the most important factors influencing task difficulty were the linguistic demands (or code complexity, as described by Skehan, 1996) of a task. The study found both Skehan’s LAC model and Robinson’s CH to partially account for the data, but also incomplete to reflect all the themes and categories revealed in the qualitative analyses. In this sense, the researchers claimed that Robinson’s triadic model fails to account for issues such as clarity of pictures, linguistic demands and story. The study concluded that data supported Skehan (1998) proposal over Robinson’s as it reflected the Cognitive component as well as the Code Complexity. One limitation attributed to Skehan’s (1998) model was that affective factors and learner-related variables (e.g. age, gender, cultural background) were not accounted for in his model.

Another study, Kim et al. (2015) examined cognitive task complexity via a stimulated recall, a subsequent interview and a task difficulty scale to validate the assumption of task complexity. Reasoning demands (+/- reasoning) was the feature manipulated for task complexity in an information exchange (simple task) and the decision-making task (complex task). In the stimulated recall session, participants were asked to watch a video of themselves while interacting with their peers and describe their mental states and thoughts. Interestingly, results revealed that, although learners rated the two tasks equally difficult, the decision-making task (the designed-to-be complex task) required more steps, analyses and evaluations as compared to the other task, and therefore, was perceived as more cognitively demanding. Results for the interview showed that task topic, the type of language required to use, and cognitive processes were perceived as the
factors affecting task difficulty.

Finally, Baralt (2013) employed time estimation under the retrospective approach. In this case, participants were required to use the retrospective approach and estimate the time they thought was spent on the task after task completion. As predicted by the researcher, findings showed that estimated time on the more complex task was significantly longer than their actual time spent on the task. On the contrary, in the simple task (where reasoning demands were lower), results indicated that the estimated time was shorter than the actual time on task.

2.9 Future Directions in Task-Based Language Teaching

Over the last five decades, SLA researchers in the TBLL domain have come a long way investigating the role of L2 tasks on learners’ task performance and L2 development with the purpose of promoting the second and foreign language learning. One of the most prominent areas of research in TBLL has been cognitive task complexity. However, due to the many different ways in which SLA researchers have operationalized this construct, scholars have recently raised concerns about the actual effects of cognitive task complexity on L2 learning (Sasayama et al., forthcoming). Specifically, researchers studying cognitive task complexity have tended to make general assumptions on the cognitive complexity imposed by tasks by simply observing the learners’ linguistic outcomes.

Recent studies (Révész et al., 2014, Révész et al., 2016, Sasayama, 2016) have addressed this issue by validating the cognitive demands of tasks using a variety of cognitive psychology measures to specifically gauge the cognitive load that L2 tasks actually impose on language learners. Therefore, to contribute to a more adequate conceptualization of cognitive task complexity, future research in SLA should incorporate
cognitive psychology methodologies (see CLT above) to examine the cognitive load that L2 tasks are actually imposing on learners. Following recent recommendations in the field (Sasayama, 2016), TBLT researchers concerned with the impact of cognitive task complexity on L2 performance should triangulate various cognitive load measures (Sweller, 1994), such as mental effort (Paas & van Merriënboer, 1994), task difficulty (Robinson, 2011), introspective measures (Kim et al., 2015), and other objective direct methodologies, such as the dual-task paradigm (DeKeyser, 1993, Révész et al., 2014) to achieve a better understanding of the cognitive demands that L2 tasks involve.

In addition, although SLA researchers, particularly in the area of TBLT, have begun to pay attention to how individual differences interfere with task performance (Robinson, 2001, 2007, 2011), the role that language anxiety plays on L2 tasks still remains today unknown. Until today, previous literature has offered an incomplete and unclear account of how language anxiety impacts learners’ L2 production concurrently to task performance. Specifically, the vast majority of studies concerned with the effects of language anxiety (Gilabert et al., 2009; Révész, 2011; Robinson, 2001b) have looked at this phenomenon under the task difficulty dimension of Robinson’s (2001a, b) TCF. However, Robinson’s task difficulty component is built upon a combination of ability (aptitude) and affective factors (motivation and language anxiety) conflated into the same single construct. In this regard, by using an approach in which individual differences are not examined in isolation, but altogether, it becomes difficult to attribute the effects observed to any of the learner’s factors studied under task difficulty. Future research needs to examine the effects of specific learner factors, such as language anxiety, in isolation from the others. A more in-depth perspective will potentially offer a clearer picture of how exactly and to what extent
this particular learner factor (instead of the combination of all of them) affects L2 task performance.
CHAPTER 3: FOREIGN LANGUAGE ANXIETY (FLA) AND L2 TASK PERFORMANCE

3.1 Conceptualization of Anxiety as a Psychological Construct: Its Origins

SLA researchers have agreed that Foreign Language Anxiety (FLA) is the affective factor with the most potential to significantly impact learners’ achievements and performance in their second language (L2) learning process (Horwitz, 2001; Baralt & Gurzynski-Weiss, 2011). However, due to the multi-faceted nature of FLA, this construct has been conceptualized and operationalized in a variety of ways, yielding conflicting results in how this phenomenon affects L2 learning (Horwitz, 2001). It is important to note previous literature on anxiety in the FL (and the L2) has generally used the terms FLA or simply language anxiety interchangeably to refer to the same phenomenon, and so will the present study.

Spielberg (1983) defined anxiety as “the subjective feeling of apprehension, tension, nervousness and worry associated with arousal of the autonomic nervous system” (p.1). Language anxiety as a concept was born from both the psychology and education fields, becoming one of the most widely investigated affective factors having an impact on L2 learning. The interest in studying language anxiety in SLA emerged from the assumption that if anxiety impaired any type of learning; it should also logically apply to language learning (Horwitz, 2001). Although several early studies (Chastain, 1975; Dulay & Burt, 1977) began to investigate the relationship between language anxiety and L2 achievement, it was not until 1986 that Horwitz, Horwitz and Cope (1986) coined the term and developed the concept of FLA as independent and unique type of anxiety characteristic to the language learning process. Specifically, these researchers highlighted that some of the reactions that language students have reported when facing the learning of a foreign
language include communication apprehension, tension, or fear of negative evaluation in the L2 among others.

3.2 Theoretical Distinctions of Anxiety as Construct: Trait vs. State Anxiety

Scovel (1978) established a fundamental categorization of the types of anxiety adopted by previous language anxiety research (Horwitz, 2001; Baralt & Gurzynski-Weiss, 2011) as an essential theoretical distinction that needs to be carefully considered when examining FLA. In particular, Spielberg (1983) has conceptualized anxiety in three main ways: (a) trait anxiety: “an inherent personality characteristic, or a more permanent predisposition to be anxious” (cited in Ellis, 2008, p. 691); (b) state anxiety: considered a temporary reaction to a particular anxiety-provoking stimulus, such as a specific test; and (c) situation-specific anxiety: generally associated with particular activities or L2 situations. Additionally, Spielberg (1966) made the following distinction between trait and state anxiety:

“Trait anxiety refers to being anxiety-prone, that is, a stable personality characteristic. Some individuals are said to be high in trait anxiety. Trait anxiety is distinguished from state anxiety, which refers to the immediate feelings of being anxious, such as nervousness and bodily tension.” (cited in Zeidner & Matthews, 2011, p. 8).

Importantly, Zeidner and Matthews (2011) further argued that the person’s state anxiety is determined by two main factors: (a) the individual’s personality traits or trait anxiety (understood as the person’s general tendencies or predisposition to feel more or less anxious) and (b) the particular external stressor (or stressful situation) that is causing those negative feelings. In addition, another important distinction needs to be made with
regards state and situation-specific anxiety. In short, the former is dependent on a stressor or anxiety-provoking stimulus at a given moment (e.g. a test), it does not, however, bear a direct relationship with a particular type of activity. However, the latter is more directly associated with a particular type of activity (such as public speaking), and therefore, it tends to be more persistent in nature. In this case, situation-specific anxiety will repeatedly and consistently be triggered whenever the individual face that type of activity. In the classroom context, for instance, if a learner “X” tends to feel anxious when he/she needs to speak the target language in front of his/her peers, the learner “X” will feel apprehension and anxiety every time he/she is required to do that task.

3.3 Foreign Language Classroom Anxiety in Second Language Acquisition

The following section reviews the most relevant studies in previous FLA literature that exist today in the SLA field.

3.3.1 Definition of Construct

As discussed in the previous section, the concept of FLA was first introduced by Horwitz et al. (1986), who argued that language anxiety stems from the inherent difficulties that learning a second language entails. Contrary to communication in the L1, in the L2 the learner necessarily faces several challenges, such as to understand and be understood, or feel evaluated by native speakers due to a lower level of proficiency in the foreign language (Horwitz, 2001). In addition to developing the concept, Horwitz et al. (1986) also created the Foreign Language Classroom Anxiety Scale (FLCAS), today’s most widely used scale that has become the standard measure to examine language anxiety in the foreign language classroom.
It is important to note that among the most prominent strands of literature on language anxiety, two of the most researched areas have been interested in: (a) examining sources that contribute to explain it, and (b) exploring the relationship between FLA and L2 achievement. The following section reviews how these two strands of research have contributed to the language anxiety literature in SLA.

### 3.3.2 Factors Associated with Foreign Language Acquisition

Among the extensive previous literature on FLA, one of the areas that has attracted more interest in the study of FLA has been that of identifying the major factors that cause language anxiety. The majority of this literature employed a variety of methods of data elicitation, such as interviews or diaries, surveys and scales or questionnaires (the latter being the most commonly employed).

Some studies (Gkonou, 2013; Hall & Graham, 2007) investigated the effects of FLA by employing diaries to record learners’ perceptions about their feelings and introspective reflections about the L2 classroom experiences. Gkonou (2013) identified several linguistic and non-linguistic factors as the major sources of classroom FLA. Interestingly, speaking was reported by learners to be the first most anxiety provoking skill, followed by listening comprehension. External factors considered to trigger language anxiety associated with the oral speaking ability were: fear of speaking in front of peers and the teacher, and the classroom setting as threatening. Among the internal and trait anxiety factors inherent to the learner were fear of making errors and difficulties in producing oral output. One of the most important contributions of Hall & Graham (2007) was its emphasis on the limitations posed by diary studies. The study strongly
recommended researchers a cautious interpretation of results when employing diaries due to the conceptual and operational difficulties related to the limited nature of diary data.

Horwitz et al. (1986) (one of the most influential studies on FLA) also shed light on factors that play a role on learners’ language anxiety in the classroom. Specifically, the FLCAS identified three major anxiety-provoking L2 situations: communication apprehension (CA), fear of negative evaluation (FNE) and test anxiety (TA) considered by the researchers as the main components of FLA.

### 3.3.3 Foreign Language Anxiety and Second Language Achievement

Today, the most prominent strand of research examining FLA consists of correlational studies that have employed anxiety scales and questionnaires to examine the relationship between FLA and L2 achievement in the foreign language classroom context (Cheng, Horwitz, & Schallert, 1999; Bailey, 1983; Horwitz, 1986; Horwitz et al., 1986; MacIntyre, 1995; MacIntyre & Gardner, 1994a, 1994b). In other words, this research domain has primarily used subjective self-report methodologies to gauge learners’ perceptions about their anxiety in the classroom setting. Overall, the vast majority of these studies have revealed that there is a negative correlation between FLA and L2 achievement (Aida, 1994; Cheng et al., 1999; Gardner, 1995; Horwitz, 1986; Horwitz, 2001; Horwitz et al., 1986; MacIntyre, 1995; MacIntyre & Gardner, 1994a, 1994b). However, a few studies found no relationship (Chastain, 1975; MacIntyre & Gardner, 1989) or even a positive relationship (Kleinmann, 1977, Spielman & Radnofsky, 2001).

Initially, a considerable number of studies examined the effects of FLA on course grades as measure of L2 achievement and found a moderate negative correlation (Aida, 1994; Horwitz, 1986; Kim, 1998; Saito & Samimy, 1996; Coulombe, 2000). Aida (1994)
showed a statistically significant negative relationship between the FLCAS and final course grades on American and second-year Japanese learners. Kim (1998) found the same negative correlation on a communicative instructional context over a traditional reading context on Korean learners of English. Coulombe (2000) and Saito and Samimy (1996) both examined L2 achievement across different proficiency levels (beginning, intermediate and advanced) on college Canadian French learners and Japanese learners respectively. Results in the two cases also revealed significant negative correlations between participants’ FLA and their final language course scores. Interestingly, Saito and Samimy (1996) found that the advanced group showed the highest levels of anxiety, followed by the beginning level, and the intermediate group yielded the lowest FLA scores.

Bailey et al. (1998) investigated this relationship on college learners across three different languages: Spanish (N=158), German (N=20) and French (N=76). As in the previous studies, participants were administered the Horwitz et al.’s (1986) FLCAS questionnaire. Even though no statistical significant differences were found between the three languages, results were in line with previous research showing a moderate negative relationship between anxiety and L2 achievement across the different languages.

Importantly, the studies reviewed above seemed to indicate a negative relationship between language anxiety and L2 achievement (as measured via course scores). In addition, this relationship was equally observed across different L2 languages (Bailey et al, 1998).

Despite the general consistent findings revealed by the vast majority of SLA literature on language anxiety, several studies (Horwitz, 2001; Horwitz et al., 1986) pointed out that relying merely on course grade was insufficient to account for L2 achievement,
since other individual factors could have explained the results obtained. These researchers emphasized the need to examine more specific L2 situations to get a better understanding of language anxiety on learners’ L2 achievement.

3.3.4 Previous Literature on Skill-Based Anxiety and Second Language Specific-Situation Anxiety

Several SLA researchers (Horwitz, 2001; MacIntyre & Gardner, 1994a, 1994b) argued that a focus on situation-specific anxiety (as opposed to a more general application to the foreign language classroom) was needed to achieve a better understanding on how learners’ individual factors affect L2 performance in a variety of different learning situations. Supporting this argument, recent literature (Dewaele & MacIntyre, 2014) have highlighted that more consistent and meaningful results had been revealed in previous FLA studies when the focus was on particular L2 situations.

Following the previous line of findings, negative correlations between FLA and L2 achievement have also been demonstrated in research concerned with specific skill-based anxiety. In this strand of literature, it is important to note that although oral anxiety (Chen et al., 1999; Mak, 2011; Woodrow, 2006) has received the most interest within the literature studying FLA; in the last decade, attention was also attracted to other linguistic skills: listening (Kim, 2000, 2005), reading (Saito et al., 1999), and writing (Chen et al., 1999; Hilleson, 1996).

One study, Pae (2012) examined the relationship between each of the four linguistic skills anxieties (speaking, listening, reading and writing) and general classroom anxiety. The study revealed that each skill-specific anxiety was independent from the others and from general FLA. Supporting Pae (2012), the literature reviewed above found positive
correlations between FLA (classroom anxiety) and skill-based anxiety, and a significant negative correlation between these two and L2 achievement. These findings were consistent with Pae (2012) in their conclusions that general classroom anxiety and skill-based anxiety were separate but related phenomena. In addition, among the most important contributions of the studies reviewed above was the introduction and development of skill-specific anxiety scales: The Foreign Language Reading Anxiety Scale (FLRAS) (Saito et al., 1999); the Foreign Language Listening Anxiety Scale (FLLAS) (Kim, 2000, 2005); Second Language Writing Apprehension Test (SLWAT) (Chen et al., 1999).

Another strand of literature has focused on the relationship between language anxiety and L2 performance on specific L2 linguistic abilities, such as vocabulary learning (MacIntyre & Gardner, 1989) different L2 production measures (MacIntyre & Gardner, 1993), and the effects of corrective feedback (Sheen, 2008). An early study, MacIntyre and Gardner (1989) found that FLA was negatively correlated with vocabulary learning in the French language via a specific French anxiety scale. Using a variety of measures, MacIntyre et al. (1993) examined how (what they referred to as) use anxiety was related to L2 performance on a composition task, a measure of French proficiency and a cloze test. Results revealed a statistically significant negative correlation between FLA and learners’ scores across all tests. Finally, Sheen (2008) examined the effects of language trait anxiety on learners’ ability to modify their oral production after being provided with corrective feedback in the form of recasts. Among the most relevant findings, post-test measures revealed that the low anxious group that received recasts produced statistically significantly higher modified output as compared to the high anxious recast group and the control group. In addition, as expected, no significant difference was found between the high anxious
recast group and the control groups. The study concluded that language anxiety is a major factor affecting learners’ ability to incorporate feedback, and benefit from it, thus having an impact on L2 production, and more generally, on L2 learning. This study supported previous evidence showing the same effects of language anxiety on recasts (Havranek, 2002), and of other individual variables (e.g. grammatical sensitivity, previous knowledge of the language, extrinsic motivation and anxiety) (DeKeyser, 1993). Importantly, these studies argued that individual differences play a role in certain learners benefitting from corrective feedback more than others.

The examples above illustrate that when the effects of language anxiety were examined in more specific L2 learning situations – such as in particular linguistic skills, other L2 performance tests (e.g. vocabulary learning, a composition task), or when pedagogical interventions are involved (e.g. recasts) – findings have proved to be more consistent.

In spite of the general agreement today about the importance for SLA research to study language anxiety in more meaningful L2 contexts (Dewaele & MacIntyre, 2014, Horwitz, 2001), many questions still remain unanswered. Furthermore, the multifaceted nature of the anxiety construct (as it is the case with most affective factors) has made it difficult for researchers today to operationalize this variable. These difficulties have led to several important limitations on how language anxiety has been captured in research designs and, therefore, the inability of establishing a direct link between the effects observed and L2 performance.
3.4 Previous Cognitive Psychology Literature and Second Language Acquisition Literature on the Effects of Anxiety on Cognitive Processing

Although most previous literature on FLA has focused on the effects of language anxiety by using broad-based measures to account for L2 achievement in instructional settings, only a few SLA studies (MacIntyre & Gardner, 1994a; Chen & Chang, 2009) have attempted to explain how language anxiety and its relationship with cognitive demands affect task performance. As a result, the role that FLA plays on learners’ cognitive demands, and subsequently on L2 outcomes, requires further investigation (Robinson, 2007).

The detrimental effects that anxiety has on cognition, and thus, on task performance have been widely demonstrated in cognitive psychology research (Derakshan & Eysenck, 2009; Eysenck, Derakshan, Santos & Calvo, 2007). Despite the evidence, the impact that language anxiety has on cognitive processes (and consequently, on linguistic outcomes in SLA) has been ignored to date.

Two decades ago, a few early studies suggested the subtle effects that language anxiety had on processing (Steinberg & Horwitz, 1986; MacIntyre & Gardner, 1989). Steinberg and Horwitz (1986) examined the effects that anxiety-arousal had on L2 performance during a picture description task. Learners were assigned to an anxious and a relaxed classroom condition (as manipulated by the researcher) to look at the effects of state anxiety. Results showed that the highly anxious participants were less interpretative when commenting on the pictures. Another study, MacIntyre and Gardner (1989) found that more anxious learners showed a slower rate of vocabulary learning and word retention.
These findings are in line with the cognitive psychology research evidencing that cognitive resources are affected by language anxiety.

MacIntyre and Gardner (1994a) investigated the effects of language anxiety on specific task performance and on the cognitive activity preceding that performance. Researchers operationalized language anxiety by applying Tobias’ (1986) three-stage model of learning, namely, input, processing, and output, on 97 first-semester French learners. The model was used to isolate language anxiety at these three levels of processing. Performance was measured via course grades and nine oral and written specific tasks especially designed to activate the cognitive processes typical from each of the three learning stages. For the input stage, recognition tasks (word span, digit span, and t-scope) were used; for the processing stage, a French achievement test, paragraph translation and pair associates, and tasks that require recalling, organization, storage and assimilation of information were included; and for the output stage, the researchers employed production tasks (thing category, cloze test, and self-description), which were analyzed for quantity and quality of speech production measures. For the particular purpose of the study, a scale was created to account for levels of anxiety during the input, processing and output stages (referred to in the study as “computerized 16-point anxometer scale”). Participants were assigned to perform the tasks in a randomized order and administered several anxiety scales: (a) the FLCAS (Horwitz et al., 1986), (b) the French Class Anxiety scale (Gardner, 1985), (c) and the French Use Anxiety scale (MacIntyre & Gardner, 1988). Results showed statistically significant negative correlations between course grades and scores for anxiety at each of the processing learning levels. Negative correlations were also found between specific input and output task performance and anxiety during those stages. The t-scope
task and processing tasks yielded a positive correlation between anxiety and latency in word recognition and reaction times. The study claimed to support Eysenck (1979) suggesting that anxiety negatively affects cognitive processing more than just in the production (or output) stage.

A more recent study (Chen & Chang, 2009) has examined how anxiety impacts cognitive load, and in turn, L2 performance in a listening task on 88 English learners in Taiwan. The argument for the researchers to investigate cognitive load is that it has been associated with working memory capacity during task performance. Participants’ cognitive load was tapped and estimated via a self-report measure, the Cognitive Load Subject Rating Scale (CLSRS). Participants were required to self-estimate the level of task cognitive load. Interestingly, researchers found that L2 learners with high language anxiety also incurred in higher cognitive load. Furthermore, results revealed that classroom language anxiety (as measured by the FLCAS) and cognitive load were in negative correlation with L2 task performance.

The studies reviewed above employed different terminology—cognitive load (Chen & Chan, 2009), and cognitive processing (MacIntyre & Gardner, 1994a)—to refer to essentially the same concept, namely, cognitive demands or resources.

3.5 Limitations in Previous Foreign Language Anxiety Literature

The following section summarizes critical gaps present in previous FLA literature in relation to how SLA researchers have so far conceptualized and operationalized the construct of language anxiety.

3.5.1 Tendency to Operationalize Foreign Language Anxiety as a Static Phenomenon

One of the main current concerns related most language anxiety research is the
methodological problem of administering anxiety questionnaires at particular times and, afterwards, accounting for performance at different times (Baralt & Gurzynski-Weiss, 2011). In other words, if the timing of administration of the questionnaires to the learners is not concurrent to when the task is actually performed, logically, language anxiety cannot be attributed to that particular L2 situation.

In this regard, despite the multiple attempts to pay attention to specific L2 contexts, SLA researchers have consistently operationalized anxiety as a static construct, without considering the immediate nature thereof (viewed as state anxiety) (Zeidner & Matthews). Specifically, most previous literature on FLA, even when focusing on particular L2 situations, have exclusively employed classroom anxiety scales designed to account for the general learner’s anxiety tendencies and predispositions applicable to the language classroom (instead of using concurrent measures applicable to the specific task being explored), thus reflecting this tendency of treating anxiety as an invariable phenomenon.

Only a few researchers to date (Baralt & Gurzynski-Weiss, 2011; MacIntyre & Gardner, 1994b) have attempted to capture state anxiety concurrent to task performance. Specifically, Baralt and Gurzynski-Weiss (2011) investigated the relationship between computer-mediated communication (CMC) and language anxiety on twenty-six English learners of Spanish during task-based interaction. Language anxiety was conceptualized in the study as the consequence to L2 task performance. In particular, the researchers examined whether carrying out an information-gap task in different types of medium - CMC and Face-to-Face (FTF) communication - would differentially affect learners’ levels of state anxiety. Whereas participants in the FTF group interacted orally, in CMC they did so in writing via an iChat. Interestingly, a questionnaire was developed for the purpose of
the study to account for state anxiety as a response to the particular task and type of medium. In addition, the study operationalized state anxiety as a temporary phenomenon that needed to be measured at several points in time throughout the task, due to its fluctuating nature, and thus reflect learners’ anxiety as a response to a particular L2 situation. Therefore, learners were provided with the same anxiety questionnaire exactly halfway through the task (during the process), and immediately after task completion (after the process). Participants were also administered a task-preference questionnaire to examine their perceptions on the tasks in both medium types. Contrary to their initial hypothesis, results showed no significant differences between CMC and FTF mode in their levels of anxiety in L2 Spanish learners.

Another study, MacIntyre and Gardner (1994b) investigated the effects of state anxiety at the input, processing and output learning states of a vocabulary learning task on college French learners. Interestingly, to examine the effects of state anxiety, a video camera was introduced to arouse anxiety at different points through the performance of the task. In particular, participants were assigned to four different groups, and for each group anxiety was aroused at different stages. For one group, anxiety arousal occurred during the stimuli presentation stage; for the second group, at the vocabulary-learning phase; for a third group at the production stage; and finally, a control group that did not have anxiety aroused during task performance. Interestingly, results revealed important increases in state anxiety for all treatment groups (at the different learning stages). The study concluded that state anxiety, as a result of a threatening situation, has serious implications for L2 learners.

Importantly, the examples reviewed above demonstrate that language anxiety is in fact a multidimensional and dynamic construct, and that it should be viewed and
operationalized as both the cause and the consequence of language learning and performance (rather than just one or the other).

3.5.2 Is Foreign Language Anxiety the Cause or the Consequence of Second Language Performance?

An issue of debate widely discussed among SLA researchers (stemming from the aforementioned static conceptualization) is the challenge to determine the directionality of the relationship between language anxiety and L2 achievement. Arguably, language anxiety can be seen as the reason for poor L2 performance, but also as a reaction (or result) to that performance (Baralt & Gurzynski-Weiss, 2011). In this respect, determining whether language anxiety should be considered as the cause or the consequence of L2 performance and achievement has raised concerns in previous literature.

Importantly, in most of the correlational studies on classroom anxiety (Horwitz, 2001; Horwitz et al., 1986), FLA has been considered as a stable phenomenon and the cause of poor L2 performance and achievement. This conceptualization has become problematic, since although it has been shown that anxiety can significantly impact learners’ linguistic outcomes (Horwitz, 2001; Horwitz et al., 1986; Sheen, 2008) (as the reason for learners’ L2 achievement), there is also evidence that it can be the result or reaction to task performance (Baralt & Gurzynski-Weiss, 2011).

The literature reviewed above has shown that most previous FLA studies have viewed FLA as either the cause or the consequence of the learning process. However, the majority of these proposals seem to offer a skewed and limited view of language anxiety that evidence important conceptual and operationalization limitations.
These challenges have been recently pointed out by SLA researchers (e.g. Dewaele, 2012; Dörnyei & Ryan, 2015) that have discussed the complexity and dynamicity of language anxiety as a construct. As part of a late influential paradigm shift, these researchers have supported the idea that language anxiety and other affective factors need to be conceptualized by treating anxiety as a ‘situated and process-oriented’ characteristic (Dewaele, 2012, p. 43). For decades, language anxiety has been considered to be in a unidirectional relationship where emotions are the causal element and the L2 outcomes are the result of that one-way relationship. However, the current dissertation supports this dynamic movement arguing precisely that language anxiety and other emotional factors cannot be viewed as either the cause of the consequence of L2 achievement or performance. Not only, it is likely that they will probably operate in both directions (rather than constituting either the source or the result of a unidirectional process), but it is believed that they should be studied as part of a dynamic system (Gkonou, Daubney, & Dewaele, 2016; Dörnyei & Ryan, 2015).

To address this limitation, these researchers based their claims on the Complex Dynamic Systems Theory (CDST) (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008) to explain the role of language anxiety in L2 tasks. The CDST attempts to captures the relationship between emotional states, cognitive and performance factors within the learner in a more complete way.

Additionally, the dynamic relationship between feeling and thinking has drawn a great deal of attention by recent studies from cognitive psychology and social psychology (Storbeck & Clore, 2007; Parrot, 2007; Golombek & Doran, 2014). In a recent volume edited by Mercer and Kostoulas (2018), new theoretical dynamic frameworks have
proposed the study of emotion as “an integral component of a three-part model consisting of emotion, cognition and behavior” (p. 144). On teacher emotions, King and Ng (2018) argued that emotions are better captured using this three-part dynamic framework (see Figure 1) that proposes a bidirectional flow between the three components. This model was based on a teacher-student model proposed by Schutz et al. (2006) on how emotion affects teaching behavior. Below, King and Ng’s (2018) model is presented:

**Figure 1**: Ng and King’s (2018) Multi-dynamic view proposed for researching teacher emotion (cited in Mercer & Kostoulas, 2018, p. 145)

Adapted from King and Ng (2018), the current study proposes a simpler version of the above model (see figure 2) to explain the relationship between language anxiety,
cognition and task performance within the L2 learner. Through this reduced version of Ng and King’s (2018) model, the current dissertation argues that the three-way interface Language Anxiety – Cognition – L2 Performance is better captured by employing a dynamic model—based on the CDS (Larsen-Freeman & Cameron, 2008)—that explains the behavior of these three components within the learner, as well as the continuous and constant interaction among them. As can be seen in the model above, the social and institutional dimensions of Ng and King’s (2018) proposal are not considered, as they are out of the scope of the current dissertation.

Figure 2: Donate’s (2018) adapted version of King and Ng’s model for the L2 learner: Interconnectedness between Language Anxiety — Cognition — L2 Performance.
In connection to the *trait - state* dimensionality, the following needs emerge and still remain open for further investigation: (a) the need to examine state anxiety (as a result of task performance) in a wider variety of language learning contexts; (b) the need to explore state anxiety as a response to different L2 tasks and task conditions; and (c) the need to investigate a more inclusive conceptualization of FLA in which anxiety is viewed as a dynamic phenomenon and both the cause and the consequence of L2 performance.

### 3.5.3 Towards a More Inclusive View of Foreign Language Anxiety

As discussed in the previous sections, research on language anxiety has shown certain limitations particularly in how the construct has been conceptualized and operationalized (Baralt & Gurzynski-Weiss, 2011; Horwitz, 2001). Traditionally, researchers have employed general measures of classroom anxiety, such as Horwitz’s (1986) FLCAS to account for situation-specific anxiety.

Importantly, Zeidner and Matthews (2011) noted the importance of distinguishing between that relatively stable personality trait in the language classroom (trait anxiety) and transitory emotional states (state anxiety). However, the scholars also argued that in order to understand anxiety in context, it is necessary to consider the two types of anxiety together, trait and state. Funder and Ozer (1983) claimed that “according to personality theories, individual differences in emotional (and indeed, all) responses vary across situations. Thus, even if a person is generally anxiety prone, there will likely be types of threats that to which they are especially sensitive and others that do not bother them very much” (p. 56-57). Supporting this argument, psychologists (Endler & Kocovski, 2001) described the multiple dimensions of anxiety via an interactional model of anxiety. Specifically, this model argues that “anxiety is a function of the dynamic interactions
between person and situation variables” (cited in Zeidner & Matthews, 2011, p 57). Furthermore, a two-way relationship exists in which human beings possess affective characteristics (more static and inherent to their personality) that also necessarily have a direct impact on how we behave in particular situations; and conversely, these specific and temporary emotional states also depend on our affective personality traits.

Applying Funder and Oxer (1983) and Endler and Kocovski’s (2001) argument to the language-learning context, a learner can have a general predisposition to feel confident and relaxed in the classroom, and yet experience high anxiety reactions when facing specific tasks at a particular moment. For example, performing an oral task in front of other learners can make a student feel apprehension, even when he or she does not tend to feel that way in the language classroom, and will in turn, be a determining factor in L2 achievement.

Spielberg’s (1966) *trait - state* anxiety theory (1966) conceptualized state anxiety as a result of an interaction between a stressful situation or event and the person’s predisposition to feel anxiety. This interaction is displayed visually in Figure 3.

**Figure 3**: Spielberg’s model of state anxiety (1966) depicting transactional *trait - state* conceptions of anxiety (cited in Zeidner & Matthews, 2011, p. 80)
Spielberg’s theoretical model shows a cyclical relationship between state anxiety and the several of components—situational and learner-related—that intervene in this relationship and with which it is connected.

The relationship between state and trait anxiety is essential for a more complete understanding of language anxiety during task performance. As discussed by Zeidner and Matthews (2011), the main reason to take into account both the trait and state types of anxiety is that state anxiety depends fundamentally on two factors: learners’ trait anxiety (or predispositions for anxiety), and the particular stressor or stimulus learners have to face at that particular situation. In this regard, although the first cannot be influenced in any way, the second can be pedagogically manipulated, since instructors can indeed play an important role in design task.

In SLA research, a more inclusive view of anxiety—in which the interconnection between trait\(^2\) and state anxiety is explored—is necessary to reach a better understanding of language anxiety and its nature as an emotional factor. In this regard, it is important to note that trait and state anxiety will be viewed as part of a dynamic system where individual characteristics, learners’ trait anxiety, interact with task features and emotional states in a dynamic way leading to a variety of L2 situations.

Ultimately, the areas of task-based language teaching (TBLT) and task-based language teaching (TBLL) could benefit strongly from a dialog with the psychology field in the study of language anxiety. Language anxiety, viewed as state anxiety, will likely vary across different L2 tasks, task conditions or instructional context, and, simultaneously

\(^2\) In the current study, trait anxiety is operationalized as foreign language classroom anxiety since it has been contextualized to a classroom setting.
depend on the particular type of language learner and his/her proneness to anxiety in the language classroom. The next section presents previous literature investigating individual differences in TBLT.

3.6 Task Complexity and Learners’ Factors in Task-Based Language Teaching

Recently, TBLT researchers (Albert, 2011; Révész, 2011; Robinson, 2011) have shown interest in how task complexity, “a stable and invariant feature of the task” (Albert, 2011, p. 241), interacts with task difficulty or “the role of learners’ factors in task performance”. In other words, how difficult learners perceive a task to be is associated with their own individual characteristics (Robinson, 2007). Specifically, Robinson (2007) argued that individual differences should play a bigger role in more complex tasks over the simple ones, and that consequently, complex tasks will lead to variation in L2 performance and development. In addition, Robinson (2011) highlighted the importance of individual differences in one of the five theoretical claims proposed in the CH, which predicts that L2 performance will change across task difficulty or learners’ individual factors.

In this regard, although researchers have mentioned the important role of individual differences in task complexity (Dörnyei, 2005; Dörnyei & Kormos, 2000; Dörnyei & Ryan, 2015; Robinson, 2001, 2011), these variables have generally been confounded as learner’s factors (all together) and examined only partially under a single construct, task difficulty.

It has been only in the past two decades that TBLT and SLA research began to show interest in how individual differences in isolation interfere with L2 performance and development. In particular, although a few studies have looked at the relationship between task complexity and certain learner factors, for example, creativity (Albert, 2011), working memory (Kormos & Trebits, 2011; Révész, 2011), and motivation (Dörnyei, 2005;
Dörnyei & Kormos, 2000), very little research exists on how affective factors influence L2 performance in TBLT.

In addition, today a current challenge in how learners’ factors have been conceptualized and operationalized in previous literature has to do with the fundamental differences associated with their particular nature. In this sense, Albert (2011) argued that certain learners’ characteristics (such as working memory capacity, language aptitude, and intelligence) have been considered as more static personality traits. On the contrary, the researcher claimed motivation or affective factors, such as foreign language anxiety (FLA), have generally been viewed as more temporary phenomena. The rationale for this argument has been attributed to the variable and changing nature of affect (or emotional variables), as opposed to the more stable nature of working memory, intelligence, aptitude or other personality-related characteristics. The following section will review the limitations observed in task complexity literature that has attempted to investigate language anxiety.

3.7 Limitations in Cognitive Task Complexity Literature on Language Anxiety

The following section discusses the current limitations found in previous research that has attempted to investigate language anxiety using the TBLT framework as main theoretical approach. These main issues observed in previous task complexity studies are related to the way language anxiety has been conceptualized, which also led to methodological problems. These limitations served as motivation for the present study.

3.7.1 Language Anxiety Framed Under Robinson’s (2001) Task Difficulty Component in Task-Based Language Teaching

Robinson (2001b) found that task role made a difference on the rating of task difficulty on 44 Japanese English learners. Task role was operationalized as the more-or-
less active role that a learner takes leading to differential effects on L2 production. Participants were enrolled in their third and fourth year of college and had received six years of previous instruction in English. In the study, they were administered the *Affective Perception Questionnaire* that assessed learners’ level of motivation and anxiety after engaging in a direction-giving map task. Among the most relevant findings were more complex tasks were perceived as more difficult, and also that task role made a difference in learners’ perceptions of the difficulty of the task. Particularly, when increasing the complexity of the task, speakers (information provider) felt more stressed and less confident than hearers (information receivers) on task performance. The more complex task led to more accurate and varied lexically language production, but also less complex and fluent for the speaker role. Additionally, task complexity enhanced more interaction on the hearer part for confirmation checks and clarification requests. The study concluded that a strong relationship existed between task difficulty, task complexity and task production (Robinson, 2001b).

Another study, Gilabert, Baron and Llanes, (2009) used the *Affective Perception Questionnaire*, adopted from Robinson (2001b), to examine learners’ perception of task difficulty, stress and confidence. In their study, participants rated the more complex task as significantly more difficult and felt less confident about completing the task.

Although these studies have revealed a significant interaction between task complexity and learners’ affective factors on L2 production and performance, there is also evidence (Révész, 2011) showing no apparent interaction between these two variables.

For example, Révész (2011) examined the effects of task complexity on form-meaning connections over an argumentative task and focused on speech production
measures. The mediating effects of individual differences, one of them being language anxiety, was investigated in the more and less complex task conditions to identify any potential interaction. Whereas results showed statistical linguistic gains for accuracy and lexical diversity as task complexity increased, language anxiety did not seem to impact learners’ speech production or, importantly, the amount and type of interaction-driven language learning opportunities. One explanation provided for the lack of effects is the advanced proficiency level of participants that might have influenced their performance. Interestingly, Révész (2011) suggested that even the most anxious learners might have developed coping strategies that explain why no linguistic differences were found.

The literature reviewed above seems to indicate mixed results regarding the impact of language anxiety as framed under the task difficulty component on L2 performance. Whereas some studies (Gilabert et al., 2009; Robinson, 2001) found that FLA influenced learners’ perceptions of how difficult more cognitive complex tasks were (versus simple tasks), the opposite findings were also revealed within TBLT research (Révész, 2011).

3.7.2 Previous Literature on Task Complexity and Language Anxiety

Surprisingly, to date only two studies (Robinson, 2007; Kim & Tracy-Ventura, 2011) have attempted to shed light on language anxiety interferes with task cognitive task (being anxiety the main variable of interest), and how this potential interaction may affect L2 task performance.

Robinson (2007) was the first study to examine the relationship between task complexity and language anxiety on Japanese learners of English, and how this interaction influenced L2 oral production. To test the Cognition Hypothesis, 40 learners were randomly assigned to three interactive narrative tasks (low, mid and high complex).
Following previous studies, the more complex tasks yielded more interaction, uptake and complex speech over the mid-complex and simple tasks. Task difficulty was operationalized to account for participants’ language anxiety via a task difficulty questionnaire administered right after task completion. Participants’ levels of anxiety were assessed via the “input, processing and output anxiety scale” (IPOAS), adopted from MacIntyre and Gardner (1994). MacIntyre and Gardner (1994) developed this scale with the purpose to specifically address language anxiety during the input, processing and output stages. Interestingly, the study revealed that the more complex the task was, the more difficult and more stressful it was perceived to be when compared to the simpler ones. Additionally, the IPOAS indicated that the levels of output anxiety (focusing of production) were higher than those during the input or processing stages in the more complex tasks over the mid and simple tasks. The researchers argued that the language instruction these learners received was more focused on passive receptive skills (listening and reading) and less on active production skills (speaking) due to the non-communicative approach of the instructional setting. In other words, since oral production was less practiced in class, logically language anxiety was higher during the output stage since learners were not familiar with this skill in class. In terms of linguistic gains, output anxiety was negatively correlated with the complexity of the language produced. The study supported the CH prediction with the low anxiety group producing more complex language in the more complex task than the high anxiety group (Robinson, 2007).

Another study, Kim and Tracy-Ventura (2011) examined the interaction between language anxiety and task complexity on L2 development of the English past tense. It was the first study that focused on L2 development with a pre-test, post-test and delayed post-
test design. Participants were 128 English learners from a beginning to a high intermediate range (as established by the TOEIC bridge test). They were divided into a high and a low anxiety group as determined by the scores obtained in a classroom anxiety questionnaire. The study reported that complex tasks enhanced L2 past tense development over the mid and simple tasks, and overall, low anxiety learners showed linguistic gains over the high anxious learners in all the tasks. The researchers concluded that FLA not only seems to impact learners’ L2 achievement negatively, but it seems to have a detrimental long-term effect on L2 development of particular difficult linguistic structures.

In short, while Robinson (2007) showed that more complex tasks led to higher anxiety levels during task performance; Kim and Tracy-Ventura (2011) demonstrated language anxiety negatively interacts with performance, and in the long run, it also has a detrimental effect on L2 development. These findings revealed the importance of taking into account the impact of language anxiety in the design of pedagogic materials and tasks. Critically, the studies also reveal major clear limitations in both the conceptualization and operationalization of language anxiety as associated with task complexity that will be highlighted in the following section.

3.7.3 Operationalizing Language Anxiety in Cognitive Task Complexity

As discussed above, to date, only two studies outside the task difficulty strand of research (Kim & Tracy-Ventura, 2011; Robinson, 2007) have specifically addressed the relationship between FLA and L2 performance in cognitively more-and-less complex tasks. However, as in the task difficulty literature, they also reveal important methodological limitations, providing an incomplete account of the role of anxiety in L2 tasks. Critically, since language anxiety is a multifaceted phenomenon that constantly
varies according to a specific L2 situation, the state anxiety dimension (viewed as a temporary online reaction triggered by an L2 specific event) needs to be carefully considered when accounting for its effects as a result of certain L2 task features, and task conditions. In this sense, Kim and Tracy-Ventura (2011) failed to capture the temporary nature of learners’ anxiety states as happening concurrently to task performance (Baralt & Gurzynski-Weiss, 2011). Instead, the researchers examined anxiety only as a static emotional phenomenon applicable to the foreign classroom (tapped by the standard measure of foreign language classroom anxiety, namely, the FLCAS, Horwitz et al. 1986).

In the case of Robinson’s (2007) language anxiety’s conceptualization and measurement has also turned out to be problematic. Specifically, the study used the input, processing, output anxiety scale (IPOAS) (adopted from MacIntyre & Gardner, 1994) to gauge language anxiety in different task complexity conditions. However, this instrument was originally designed to investigate learners’ perceptions of their regular levels of anxiety experienced during comprehension, processing and production tasks in the language classroom (rather than specifically using a state anxiety measure). In this case, it was the temporary dimension of language anxiety (as a result of task performance) that was not addressed.

In addition, although cognitive psychologists (Derakshan & Eysenck, 2009; Eysenck, Derakshan, Santos & Calvo, 2007) have demonstrated the detrimental effects that anxiety has on cognitive processes (especially on attentional control processes), surprisingly, this relationship has barely been investigated in the SLA field (MacIntyre & Gardner, 1994). Chen and Chang (2009) and MacIntyre and Gardner (1994) found that classroom language anxiety was negatively correlated with L2 performance, and that these
effects were explained by how anxiety, in turn, affects cognitive processing. In short, both studies suggested that low anxious learners outperformed high anxious learners mainly due to the slowing mediating effects that anxiety has on cognitive processing. Despite this evidence, no researcher has yet attempted to specifically examine this relationship between anxiety and cognition in TBLT research.

Although many previous studies have shown the linguistic benefits that a cognitively more complex task has on L2 task performance (Robinson, 2011), the moderating role of language anxiety in L2 tasks has still to be investigated. Arguably, more cognitively complex tasks (versus more simple ones), by imposing higher cognitive demands, may potentially affect high anxious and low anxious learners differently in their L2 production during task performance. Additionally, how learners perceive their state anxiety levels during cognitively demanding tasks (as compared to less demanding tasks), and, in turn, how anxiety seem to affect L2 linguistic outcomes is today unexplored in the area of TBLT.

Arguably, learners that are less predisposed to feel anxiety in the foreign language classroom (low anxious learners) may naturally also show lower state anxiety levels in a cognitively more complex task (versus a simple task) as compared to high anxious learners. Furthermore, more complex tasks may logically trigger higher levels of state anxiety over the simpler tasks. In this respect, considering the relationship between learners’ proneness to anxiety in the L2 classroom and state anxiety as taking place during task performance may be useful to get a better understanding of the role of anxiety in L2 tasks.

Validation of the actual cognitive demands of L2 tasks by examining their cognitive load (Sweller, 1994) via independent measures of cognitive complexity is essential to
understand how state anxiety may be connected to cognitive processing in the L2 classroom. A potential interaction of cognitive load measures and state anxiety has not been investigated in the same study to date, specifically in L2 Spanish.

Increasing cognitive demands in classroom pedagogical activities may (or may not) affect Spanish learners differently depending on their individual characteristics. In addition, language anxiety has been shown to be both facilitative and detrimental. An insight into both the cognitive and emotional (particularly anxiety) processes would provide a more complete picture of what task characteristics better promote L2 learning.

3.8 Summary of the Chapter

To summarize, while there is a significant amount of literature in SLA and TBLT that has investigated the effects of language anxiety on L2 performance from different perspectives, the vast majority of research have been focused on correlational studies. Such research designs present important limitations and offer an incomplete picture of what language anxiety is and its complicated behavior in the L2 instructional setting. Despite the existing evidence of the negative impact of this affective factor on L2 development and learning, the complexity of language anxiety as a construct, its multiple dimensions and facets and has posed considerable challenges in its conceptualization and operationalization in research designs, leading to mixed and inconsistent findings today.

Recent observations made by SLA researchers (Baralt & Gurzynski-Weiss, 2011; Gkonou, Daubney & Dewaele, 2016; MacIntyre & Serroul, 2015) are mainly related to language anxiety treated as a static phenomenon, and thus, ignoring its dynamic and continuously changing nature. Much of the prior research has employed instruments that present methodological problems by examining general classroom anxiety rather than
state anxiety as an emotional state particular to a concrete L2 event. Most of these research designs have attributed the language anxiety to a particular activity or task, but instead captured it at a different point in time during the experimental process. The present dissertation intends to address the dynamic nature of language anxiety by examining the *trait-state* dimensionality (Scovel, 1966), as well as the inseparable relationship between cognition, affect and L2 performance (Waninge, 2015).
CHAPTER 4: THE CURRENT STUDY

4.1 Purpose of the Current Study

Based on cognitive psychology research (Zeidner & Matthews 2011) supporting the multiple dimensions of anxiety (both trait and state anxiety are two sides of the same coin), a more unified conceptualization of language anxiety is proposed. The role that state anxiety (or immediate emotional states as a response to a particular task) plays during task performance when cognitive task complexity is manipulated was explored in the current study in relation to learners’ language anxiety predispositions towards the Spanish language classroom. In this regard, foreign language classroom anxiety (proneness to anxiety or anxiety predispositions in the Spanish classroom) was viewed as baseline for state anxiety.

In addition, the relationship between FLA, cognition and L2 performance was examined following a task-based approach that focuses on how more-and-less cognitively complex tasks are related to the learner’s state anxiety to address the limitations observed above. Following recent studies (e.g. Sasayama, 2016), the construct of task complexity was addressed under a cognitive psychology theoretical approach—Sweller’s (1994) CLT—that taps the cognitive load or processing demands imposed by tasks to validate their supposedly cognitive complexity. Supporting this view, rather than assuming the complexity of the task (by focusing merely on linguistic differences) a combination of different cognitive load measures was used to gauge cognitive task complexity. This approach not only ensures a robust operationalization of cognitive task complexity, but

---

3 Following Sasayama (2016), cognitive load is used in the present study as the main theoretical construct to validate the cognitive demands imposed by a pedagogic task in the foreign language.
also allows for an examination of the unexplored relationship between perceived state anxiety, cognitive processes and L2 task performance.

To achieve a better understanding of this relationship, the three-way interface *Anxiety – Cognition – L2 Performance* was investigated in relation to two levels of cognitive task complexity. To this end, the possible following relationships are examined:

a) The relationship between cognitive task complexity and cognitive load—as assessed by task difficulty, mental effort, and time estimation;

b) The relationship between cognitive task complexity and state anxiety concurrently to task performance;

   ➢ The relationship between foreign language classroom anxiety (used as baseline for task anxiety scores) in L2 Spanish and state anxiety during a more cognitively complex task versus a cognitively simple task was examined.

c) The relationship between state anxiety (as perceived by learners) and task performance in L2 Spanish (as measured by the CAF triad); and

d) The relationship between cognitive task complexity and L2 performance in Spanish

4.2 Research Questions

As reflected in the previous relationships, the present study attempts to contribute to the TBLT domain by shedding light on: (1) the role of cognitive load—as measured by subjective measures: task difficulty, mental effort and time estimation—in cognitive task complexity in L2 Spanish; (2) the role that state anxiety (as perceived by learners) plays in Spanish L2 tasks, and a better understanding of how cognitive task complexity is associated
with levels of state anxiety during task performance in Spanish; (3) a fuller understanding of the relationship between language state anxiety and L2 performance (as measured by the CAF triad); and lastly, (4) to tease apart the relationship between cognitive task complexity and task performance in L2 Spanish. The above partial one-way relationships were examined separately to gain a better insight on the three-way interface *Language Anxiety—Cognition—L2 Performance* in an instructed second and foreign language acquisition context. In view of these goals, the following research questions were posed:

**RQ1.** How does presumed cognitive task complexity relate to cognitive load in L2 Spanish, as measured by *task difficulty, mental effort, and time estimation*?

**RQ2a.** How does presumed cognitive task complexity relate to language state anxiety—as perceived by learners—during oral task performance in L2 Spanish?

**RQ2b.** Is there any relationship between language state anxiety—concurrent to oral task performance—and foreign language classroom anxiety in Spanish?

**RQ3.** Does language state anxiety—as perceived by learners—fluctuate over time when presumed cognitive task complexity is manipulated during oral task performance in L2 Spanish?

**RQ4.** In an examination of the three-way interface *Language Anxiety—Cognition—Task Performance* in L2 tasks.

**RQ4a.** How does language state anxiety relate to cognitive processes—as perceived by learners—during oral task performance in Spanish when cognitive task complexity is manipulated?
**RQ4b.** How does language state anxiety relate to L2 performance—as perceived by learners— during oral task performance in Spanish when cognitive task complexity is manipulated?

**RQ5.** How does presumed cognitive task complexity relate to learners’ task performance in L2 Spanish?

### 4.3 Methodology

In what follows the method employed in the present dissertation will be explained in detail. This includes participants, instruments and materials, main procedures, coding and scoring, and lastly, data analyses.

#### 4.3.1 Pilot Study

A pilot study on a much smaller scale was conducted in Fall 2016 prior to the main data collection. The purpose of the pilot study was to test the materials, and to see what worked and what did not, and thus, had to be modified for the main study. Especially important in this exploratory phase was to examine the main two tasks in relation to its potential relationship with language state anxiety and how participants reacted to them in this regard.

At that stage of the process, as for Sasayama (2015, 2016) the pilot study allowed the researcher to:

1) Validate that the intended differences on the actual cognitive demands imposed by the two tasks was indeed statistically significant different. This validation was essential to investigate the unifying thread of the current dissertation: the relationship between language anxiety, cognition, and L2 task performance.
2) For the pilot study, the following cognitive load measures were used to examine the relationship between cognitive task complexity and cognitive load in Spanish as an L2 and FL. In this regard, testing if this relationship was still maintained in Spanish. Would help validate these measures in other foreign languages.

3) Validate the state anxiety instrument. For the purpose of the current study, the state anxiety questionnaire employed in the current dissertation was adapted from Baralt and Gurzynski-Weiss (2011). For the quantitative part of the measure, new statements had been incorporated and, some eliminated, and others adjusted. Therefore, validating the instrument was of upmost importance prior to data collecting for the main study.

For the pilot study, thirty-six participants participated in the pilot study. They were selected from three different levels of L2 proficiency Spanish according to their enrollment in the language courses offered at a non-intensive program in a Northeastern University. In particular, learners had been recruited from their second, fourth and sixth semester along in the program, which corresponded to an introductory 2, intermediate 2 and advanced 2 courses at this institution in particular. Selecting participants at these three “presumably different” levels of proficiency helped the PI to be able to examine whether differences existed in the levels of state anxiety for the simple and the complex tasks across different proficiency levels.

Collecting data on the three proficiency levels served to justify a decision on what level would be more adequate for the two main tasks. Findings on the pilot study revealed more interesting results at the lower levels of proficiency regarding participants’ state
anxiety. Additionally, regarding the two main tasks, data in the pilot study also demonstrated that the main tasks were too easy to find meaningful differences across the advanced group. For this reason, a lower proficiency level proved to be more adequate for the main tasks employed and for the purpose of the current dissertation.

Importantly, in the pilot study the researcher had a five-minute time constraint controlling for time on task (TOT) following Sasayama’s research design. Nevertheless, findings revealed that establishing a five-minute time limit significantly increased learners’ levels of anxiety during task performance. A timer had been set on the computer screen while participants were performing the task showing the remaining time. The researcher observed the five-minute time limit significantly contributed to trigger participants’ state anxiety during the experiment. To ensure that levels of state anxiety could be mainly attributed to the cognitive complexity of the task, the time constraint was eliminated to reduce—to the maximum possible extent—any external factors that could potentially increase the levels of language anxiety during task performance.

In the pilot study, as opposed to most TBLT previous research, the time estimation measure had been administered using the prospective rather than the retrospective approach. Under this approach, prior to the task, participants were asked to estimate the time they thought they would spend narrating the story. The premise is that the judged time tends to decrease as cognitive load increases, as proposed by the cognitive psychologists Block et al. (2010). However, findings for this measure in the pilot study showed very abnormal distributions, proving the prospective approach to be ineffective, at least for this particular task conditions. In addition, comparisons to previous literature could not be made, as research on task complexity using this approach is currently non-existent.
Therefore, for comparability reasons, the researcher decided to apply to retrospective approach following recent studies (Baralt, 2013; Sasayama, 2015, 2016).

To conclude, all measures in the pilot study had been administered as paper-based questionnaires. However, for the main study, an electronic version of all self-assessed measures was designed. As highlighted above, this aided the coding process, and also minimized the risk of errors associated with data entering and data loss.

4.3.2 Current Study

4.3.2.1 Participants

Fifty-one L1 English learners of Spanish were selected for the present study on a voluntary basis. Participants were homogenous regarding their age, nationality, and their L1, and were undergraduate learners enrolled at a Northeastern US University.

Participants were recruited via an email announcing the second language experiment, where they could sign up and received one point of extra-credit for their participation. They were enrolled in their third semester of Spanish of the non-intensive track of a task-based language program. Participants regularly met three times a week for a fifty-minute class time period and were used to perform oral communicative tasks in the classroom. As pointed out above, in the pilot study, data from different proficiency levels was collected and analyzed to better understand how language state anxiety was related to cognitive complexity and L2 performance the different proficiency levels. Findings showed the two tasks—both the simple and the complex—were too easy for the advanced group, and way too difficult for the novice group. Therefore, participants enrolled in the
third semester of Spanish—the intermediate 1 level (according an institutional measure of proficiency) (Thomas, 2006)—seemed to be the most appropriate level for the purpose of the present study.

4.3.2.2 Materials

This section provides a description of 1) the two main tasks, 2) cognitive load measures (perceived task difficulty, perceived mental effort and time estimation), and 3) language anxiety self-rating measures (quantitative and qualitative questionnaires).

4.3.2.2.1 Treatment Tasks and Manipulation of Cognitive Task Complexity

Two main oral narrative tasks adopted from Sasayama (2016) were used for the main study. Tasks were monologic and participants were required to narrate a story in Spanish based on a six-framed set of pictures. However, rather than the four-level continuum of cognitive task complexity used in Sasayama (2016), the present study opted for two selecting only two levels for cognitive task complexity: the cognitively simplest—Task 1—and the cognitively most complex—Task 4. Originally, the tasks were adapted from Hill (1960) (task 1: “the alarm clock”) and Elder and Iwashita (2005) (task 4: “the dog”). The rationale behind choosing these tasks was: first and foremost consistency with previous task complexity literature, since most previous studies have employed picture-description (also called narrative tasks in TBLT) tasks; second, in Sasayama (2016) these two tasks in particular revealed the greatest differential effects in terms of L2 outcome measures and cognitive load measures; and third, the larger differences in the cognitive

---

4 L2 Proficiency was determined by institutional status, and it is therefore a limitation of the current study.
load imposed by the tasks may more likely lead to a potential interaction with language anxiety.

The main task feature was manipulated along the resource-directing line of Robinson’ (2001, 2010) approach, the [+/- few elements] dimension, following Sasayama’s (2016) task design. It was selected based on being the most commonly used manipulation by TBLT researchers to date and has been demonstrated to contribute to pose significantly higher cognitive demands on learners. This would also help maximize comparisons with previous literature. Specifically, in this study, the number of elements was manipulated as the number of characters, including only one character in the simple task versus nine characters in the complex task.

Following Sasayama (2016) regarding the task format, some of the features—the storyline, background information and the complexity of the language elicited—differed in the two tasks with the purpose of avoiding practice effects. However, certain features were maintained constant for the two tasks to ensure there was no increase of extraneous cognitive load (Sweller, 1994). In this regard, following Sweller (1994), ensuring that the cognitive load manipulated was only the facilitative type was especially taken into consideration as part of the task design. Furthermore, a different topic would also elicit different content and vocabulary avoiding task effects.

Also following Sasayama (2016), specific instructions were provided to participants that their task performance would be assessed based on its completeness (the six photos had to be included), effectiveness (the correct order of the pictures) and creativity of the story. Importantly, in the current study, it was decided to not set a time constraint for participants to tell their stories, as opposed to Sasayama (2016), where there
was a time constraint of 5 minutes for time on task (TOT). The pilot study had revealed setting a time constraint was a source to trigger participants’ anxiety during task performance. Therefore, for the present study, it was important to control for factors that could lead to induce learners’ anxiety during the experimental session other than the task complexity in itself. In addition, only 30 seconds were given prior to the story-telling so that planning time would not play a role in the subsequent task performance.

Tasks were presented in printed black and white hard copy, and each picture’s size was 5.50 x 5.80 inches so the characters and other items in the photos were visible enough. Table 2 below shows a summary of the main task features for the simple and the complex task according to Sasayama’s (2016) categorization: a) number of characters, b) storyline, c) storyline clarity, d) linguistic demands, and e) changes in story setting. Participants’ perceptions from the questionnaire’s open-ended questions were used to describe each dimension of the tasks.
### Table 2: Design of main tasks.

<table>
<thead>
<tr>
<th>Simple Task</th>
<th>Number of characters</th>
<th>Storyline</th>
<th>Storyline Clarity</th>
<th>Linguistic demands (code complexity, Skehan, 1996)</th>
<th>Story setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Task</td>
<td>1</td>
<td>Clear storyline and easy sequence of events.</td>
<td>Relatively easy vocabulary related to everyday life and house items: <em>cama</em>, <em>ventana</em>, <em>cuarto</em>, <em>mesa</em>, <em>hombre</em> (bed, window, bedroom, table, man).</td>
<td>No changes in setting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(A man as main character).</td>
<td>Examples:</td>
<td>Relatively easy grammar: mostly use of present, simple past tense, and high frequency verb forms: <em>dormir</em>, <em>ir</em>, <em>mirar</em>, <em>levantarse</em>, <em>acostarse</em> (sleeping, going, looking, getting up, going to bed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>From picture 1 to 3: A man is setting up his alarm clock before he goes to bed at nighttime. Afterwards he looks out his window, turns the light off and goes to bed. From picture 4 to 6: the alarm wakes him up in the morning while the sun comes in through the window. However, he is so upset that he throws the pillow to the alarm clock, knocking over both the table and the clock, and continues sleeping.</td>
<td>“It all took place in one room and there was only one character, so it was a more calming experience than the first task with the dog.” Participant 9: “Nothing was too hard- the pictures were clear and simple to understand, but not knowing some of the important vocabulary makes me feel more anxious about the Spanish-speaking.”</td>
<td>Low frequency words: <em>despertador</em>, <em>almohada</em>, <em>persianas</em>, <em>cortina</em>, <em>lámpara</em> (alarm clock, pillow, blinds/curtains, lamp).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Clear storyline and easy sequence of events.</td>
<td>• Low frequency verb constructions, and expressions: <em>cerrar las cortinas</em>, <em>persianas</em>, <em>lanzar la almohada</em>, <em>mirar por la ventana</em> (close the blinds/curtains, throw the pillow, look through the window).</td>
<td>Main language functions present: describing and narrating (Sasayama, 2016)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

5 The selection of the simple task was based on the assumption that the majority of vocabulary was accessible and relatively simple in the picture prompts. The low frequency word index is reported according to participants’ retrospective comments after task completion and will be discussed in the study results section.
Table 2. (Cont’d)

<table>
<thead>
<tr>
<th>Number of characters</th>
<th>Storyline</th>
<th>Storyline Clarity</th>
<th>Linguistic demands (code complexity, Skehan, 1996)</th>
<th>Story setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex Task</td>
<td>9</td>
<td>From picture 1 to 3: A couple is having coffee in the patio of a restaurant downtown. Next to the restaurant, a man is working on a ladder, and further away, another man is walking his dog. Suddenly, the dog escapes from his owner and runs into the middle of the street where a car is driving. To avoid running over the dog, the car swerves, but ends up running into the sidewalk, and hitting the man on the ladder. From picture 4 to 6: an ambulance then comes to takes the injured man to the hospital. When the police arrive at the scene, they are about to arrest the driver, but the driver insisted it was the man with the dog’s fault. In the last image, the man and the dog are put into in jail.</td>
<td>• Relatively unclear storyline, more complex sequence of events, and more detail. Example: <em>I felt more overwhelmed in this story because the photos were more complex. It was difficult to find the storyline because there was much more action and background in each photo.</em> • Confusion with the different characters. Example: <em>Since there were more characters, it was harder to keep track of. Also, I didn’t know how to differentiate between the different men.</em></td>
<td>• More complex vocabulary associated with particular events (e.g. a car accident, street elements,): <em>acera, escalera, carcel, problema, dueño, correas, esposas</em> (sidewalk, ladder, jail, trouble, owner, leash, handcuffs). • Relatively difficult grammar and higher proportion of more complex verbs: use of cause and effects expressions (e.g. verb periphrasis constructions): <em>golpear, caerse, arrestar, hacerse daño, salir corriendo, pasear al perro, dar un volantazo</em> (to hit, to fell, to arrest, to hurt oneself, to running away from, to walk a dog, to swerve).</td>
</tr>
<tr>
<td>(four men, a woman, a dog, a policeman, and two ambulance attendants)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.3.2.2 Self-Assessment Anxiety Measures

Two subjective self-rating measures were employed to examine language anxiety. First, to investigate state anxiety as concurrent to the task, participants were administered two adapted versions of the state anxiety questionnaire from Baralt and Gurzynski-Weiss (2011). The original questionnaire included 15 items. However, for the purpose of the current study, the new scale was reduced to 10 items and made specific to Spanish as a foreign language. Item 5, 13 & 14 from the original instrument were eliminated because they were considered redundant. Item 6 was also eliminated because it captured general foreign language classroom anxiety instead of state anxiety (‘I feel confident in my ability to quickly learn new things in Spanish’), and this is a different construct.

Items 2, 8, 12 and 15 from Baralt and Gurzynski-Weiss’ (2011) state anxiety measure were kept the same due to their general applicability to a variety of L2 communicative tasks. Critically, and since one of the main goals of the study was to explore the relationship between state anxiety, cognitive processing and L2 performance, two items from the original questionnaire (item 3 and 9) were rephrased to elicit this relationship from the learners. These items addressed information processing in connection with state anxiety and task performance:

Item 4: ‘I am getting flustered because I can’t process so much information and then talk about it’.

Item 9: ‘I am relaxed and comfortable processing the information of this task and then talking about it’.
Items 4 & 11 from the original instrument were rephrased to be more specific about what learners were actually processing. They were also adjusted to the narrative nature of the main tasks, which presented participants with a storyline they had to tell out loud:

Item 5. ‘I don’t have **enough time to figure out the storyline** before I have to talk.’

Item 10. ‘I feel like I have **enough time to figure out the storyline** before talking.’

Additionally, items 1, 4, 7 and 11 were also modified to reflect the task condition of the main task regarding its interactive factors. In this sense, due to the one-way monologic nature of the tasks, where no interlocutor was involved, the items were rephrased by referring “talking” or “speaking” rather than to “communicating or responding to a partner”.

Item 2. ‘I am not bothered by having to **speak** in Spanish so quickly.’ (I.’I am not bothered by my partner communicating quickly.’)

Item 4. ‘I don’t have enough time to figure out the story line before **I have to talk**.’

(I.’I feel like I don’t have enough time to think before I have to respond.’)

Item 7. ‘I feel tense having to **talk in Spanish** so quickly.’

Item 10. ‘I feel like I have enough time to figure out the storyline before **talking**.’

The scale was modified to a six-point Likert scale range, instead of five to avoid the usual methodological limitation of participants’ selecting the mid-number as the default option. Following Baralt and Gurzynski-Weiss (2011) measurement of state anxiety, participants completed the questionnaire at two points in time: halfway through the task and immediately after task completion. This was done to explore the dynamic nature of anxiety. Since state anxiety constantly fluctuates, it was interesting to look at how participants perceived their levels of anxiety to be during task performance and also after
the task is completed to see if significant differences existed over time. It could be argued that perceived language anxiety is shown at its highest point in the middle of a pedagogical task. However, it could also be expected that once the task is finished, the learner’s perceived levels of anxiety would tend to decrease. Additionally, participants were asked to respond to four open-ended questions divided up in two sections. The main goal of these questions was to collect further details on participants’ state language anxiety during their task performance. The first section consisted of two broad questions. Specifically, the first question (more exploratory) sought to find out general feelings (good or bad) towards the simple and complex task. 1.a. *Please explain in detail all the feelings—good and bad—that this task provoked in you, and the reasons why you think you felt that way while performing.*

After that, participants responded to a follow-up question about their performance: 1.b *How did that affect your performance in this task?* For the second section, participants were required to provide more specific information about what particular task dimensions and features had a negative impact on their task anxiety. To be consistent with the previous section, the researcher asked: 2.a. *Please explain in detail what elements of the following task—storyline, characters, length, grammar you used, vocabulary, etc.—made the task more or less anxiety-provoking and how it did so.* 2.b. *How did that affect your performance in this task?* As in the previous section, the second question addressed the relationship between participants’ anxiety and how they perceived their performance in Spanish was during the two tasks.

Second, Horwitz’s (1986) (FLCAS) was used to tap foreign language classroom anxiety, as it is the standard measure and the most popular instrument to examine classroom anxiety. This allowed maximizing comparisons with previous SLA literature on language
anxiety. Since this scale was not designed for a specific target language (rather for the foreign language in general), it was adapted to Spanish for the particular purpose of the study. For comparison reasons, the instrument was also adjusted to a 6-point Likert scale to keep all self-report measures in the same constant range. Scores on the FLCAS determined participants’ proneness (or predispositions) to anxiety in the Spanish classroom. In addition, they provided a baseline to explain learners’ levels of state anxiety as a response to the different manipulations of cognitive task complexity. Participants’ scores on classroom anxiety were compared to the state anxiety scores for the simple and the complex task to see if differences between the two tasks were revealed. This was done to investigate the relationship between the levels of task anxiety when cognitive task demands increase and their regular predispositions to anxiety.

The principal investigator (PI) designed all language anxiety measures using the software google forms to administer them via an electronic computerized version of the questionnaires. Google forms allowed the PI to email all subjective measures as one single electronic form that participants received via email. This software also facilitated a safer data collection for each measure and from each participant, automatically saving all responses to google drive. One of the further advantages of using google forms was an easier data retrieval afterwards for the coding and analysis process. Because participants’ responses and ratings were automatically saved, this avoided data loss and potential coding errors when entering individual scores.
4.3.2.2.3 Cognitive Task Complexity Measures: Task Difficulty, Mental Effort, and Time Estimation

Cognitive task complexity was investigated by a combination of subjective self-rating measures, namely, task difficulty, mental effort (Gilabert, 2007; Shiau & Adams, 2011) and time estimation (Baralt, 2013; Gilabert, 2007; Sasayama, 2016; Shiau & Adams, 2011), which helped with triangulation and validation of cognitive task complexity.

To examine perceived task difficulty, participants received the Post-Task Difficulty Questionnaire (adopted from Robinson, 2001). This consisted of one question on a 6-point Likert scale, 1 being ‘very easy’ and 6 ‘very difficult’, which required participants to rate their perceived difficulty on the two tasks. It is important to note that the task motivation and task achievement sections from the original version of this instrument from Robinson (2001) were eliminated, since these constructs did not pertain to the purpose of the present study. Following Sasayama (2016), in order to obtain specific information on what task features or other learners’ factors could have influenced task performance, participants answered to an open-ended question: What made the storytelling task easy/difficult? Please provide details.

For mental effort, the questionnaire followed Paas, van Merriënboer and Adam (1994). To examine the perceived amount of mental effort, participants were administered one question (b) How much mental effort did you put in to complete this task? Again, based on a 6-point Likert scale, participants had to self-assess their cognitive engagement in the two tasks, 1 being ‘very low mental effort’ and 6 ‘very high mental effort’, based on their own perception. After that, also based on Sasayama (2016) participants were asked to complete an open-ended question: What made the task require more or less mental effort?
Please provide details. The prompt required participants to justify and explicate their previous ratings on the amount of mental effort exerted. As explained in the literature review section, it is important to consider that although these two measures examine two different constructs (task difficulty and mental effort, respectively), they contribute to a fuller understanding of cognitive task complexity.

Another measure from cognitive psychology, time estimation was used in the study. Following TBLT previous research (Baralt, 2013; Sasayama, 2016), for this measure, after task completion, participants were asked to estimate the time they thought they had spent narrating the story. Specifically, participants were administered a handout with the following question in English: How long do you think your story was? Please indicate the approximate number of minutes and seconds you estimate to have spent in this task. They replied to the question by filling out the number of minutes and seconds in a blank space.

All cognitive load measures were administered right after participants finished narrating their stories. First, learners completed the task difficulty section. After that, followed the mental effort questions; and lastly, they judged the time spent on each task under the retrospective approach (e.g. Baralt, 2013). As for the state anxiety measures, the researcher employed google forms to deliver these questions. All measures were part the same electronic-based self-assessed post-task questionnaire.

4.3.2.2.4 Background Questionnaire

An initial background questionnaire collected information regarding participants’ personal data, language background information, and language skills in Spanish as a second or foreign language.
The questionnaire was divided into four sections. The first section elicited biodata such as participants’ gender, level of education, race & ethnicity, nationality. Section 2 continued by asking participants about general language background information regarding their L1 and country of origin. Section 3 asked participants whether they were born in a foreign country, their age of arrival (AOA) in the US, years living in the US, and their parents or caretakers’ language background information. This question aimed to identify them as heritage language learners or regular L2 learners. If participants were identified as being heritage learners of Spanish, the questionnaire automatically directed them to a different section. This was relevant for the purpose of the study, since it was determined that perceived state anxiety may significantly vary across these two groups.

A fourth part of questionnaire—*Non-native language background*—included questions referring to participants’ L2. Information gathered in this section was learners’ second and foreign languages spoken, years of formal instruction in the L2, language experience, and study abroad experiences in Spanish speaking countries. Finally, a last section of the questionnaire asked participants to judge their language ability in their foreign language(s) for listening, speaking, reading and writing. Participants were asked to rate their language abilities in English, Spanish, and all the additional languages they could use. The scale employed the ACTFL L2 proficiency categories for self-assessment: *novice, intermediate, advanced, superior and native.*

4.3.2.3 Procedure

For the current study, participants read and signed an IRB consent form that explained the purpose of the study, agreeing to take part in the experiment on a volunteer basis, as well as for permission to be recorded. After reviewing the consent form, the
researcher made clarifications and answered questions when necessary. First, participants completed the background questionnaire, followed by detailed instructions in electronic format about the main treatment tasks and the mid- and post-task questionnaires. To avoid test effects, rather than performing a practice task, as Sasayama (2016) did in her research design, the current study opted to eliminate the practice trial prior to task performance.

The two treatment tasks—the cognitively simple and the cognitively complex—were assigned in randomized order to avoid practice test effects. Participants were recorded by the program *QuickTime*. They had been informed that halfway through the story, they would be interrupted to complete a brief questionnaire, and that afterwards; they should simply continue telling the story until the end. When they had finished talking about picture 3, participants had to raise their hands to show the researcher they were halfway through the task. At that moment, participants were administered the perceived state anxiety questionnaire for the first time. Completing this mid-task questionnaire lasted approximately 2 or 3 minutes. After filling out the self-assessed state anxiety questionnaire, participants resumed the story from picture 4 to picture 6. After completion of the task, participants were administered the self-rating measures, in the following order: post-task perceived state anxiety questionnaire, task difficulty, the mental effort self-report measure, and the time estimation measure. For the latter, participants were requested to judge how long they considered it took them to narrate the whole story. Participants wrote down in a handout the approximate number of minutes and seconds they had estimated their story lasted.
For the post-task questionnaire, they first self-evaluated the Likert-scale sections, and afterwards, they were asked to complete the open-ended questions for their perceived state anxiety, task difficulty and mental effort. This stage lasted between 10 and 15 minutes.

After the first task, participants performed the second main task and the same process was repeated. The classroom anxiety questionnaire (FLCAS) was administered at the end of the session. The whole experiment took place in this single session that lasted approximately from minutes to an hour.

4.3.2.4 Scoring Procedures

The following section includes a full description of the scoring scheme and main analyses performed for each independent measure: the state anxiety self-rating measure, the classroom anxiety questionnaire, and the cognitive load measures (time estimation, task difficulty, and mental effort). After an in-depth review of the cognitive task complexity measures, the L2 linguistics outcome measures for fluency, accuracy and syntactic complexity and their respective units of analyses were examined in detail. For the constructs of L2 state anxiety, task difficulty and mental effort, the numerical and explanatory comments by participants (open-ended questions) were considered below separately as quantitative and qualitative measurements.

4.3.2.4.1 Independent Cognitive Load Measures: Scoring for Time Estimation, Task Difficulty and Mental Effort

For the time estimation measure, *QuickTime* was used to obtain participants’ actual time on task. Following the scoring system from previous cognitive psychology research (Block et al., 2010; Brown, 2008), and from TBLT previous studies that employed this measure (Baralt, 2012, 2013) scores for each participant’s actual time on task were
subtracted from scores of their estimated time to obtain the difference between the two. After that, mean scores were calculated for these difference scores for the cognitively simple and cognitively complex task. For instance, if a participant had spent 3 minutes on the task, and the estimated time is 5, participant’s score was 2. This process was performed for each participant’s time judgment on both the cognitively simple and the cognitively complex task.

Although task difficulty and mental effort provide information about the same phenomenon, the cognitive demands of tasks, they measure different constructs. Thus, participants’ responses on the two questions were coded on the 6-point Likert scale and scores were analyzed separately for each task.

4.3.2.4.2 Quantitative Measurements of State Anxiety

The following section includes a detailed description of the scoring procedures for:

a) the quantitative data for the two language anxiety questionnaires, the cognitive load measures and b) the qualitative data for all self-assessed measures.

4.3.2.4.2.1 Scoring for Self-Assessment Anxiety Measures

State anxiety scores for Time 1 (halfway through the task) and Time 2 (after task completion) were calculated following Baralt and Gurzynski-Weiss’ (2011) coding scheme. As discussed in the materials section, the usual 5-point Likert scale was adapted to a 6-point range to keep the Likert scale range constant with the other self-rating measures (mental effort and task difficulty). Anxiety statements (e.g. ‘This task is stressful for me’), ‘strongly agree’ received an anxiety score of 6, ‘agree’= 5, ‘slightly agree’= 4, ‘slightly disagree’= 3, ‘disagree’=2, and ‘strongly disagree’ corresponded to a level 1 of anxiety.
Reverse statements (e.g. ‘This task does not make me anxious’) or positively phrased items (e.g. ‘This task is fun and enjoyable’) were reverse coded.

Classroom FLA anxiety scores followed Horwitz et al.’s (1986) scoring of the FLCAS. For consistency purposes with the state anxiety scale, the Likert range was also adjusted to 6-point scale in this case. For regular anxiety statements (e.g. ‘I tremble when I know I am going to be called on in the Spanish class’) participants received (from highest to the lowest score) 6 points for ‘strongly agree’, 5 points for ‘agree’, 4 points for ‘slightly agree’, 3 points for ‘slightly disagree’, 2 points for ‘disagree’, and 1 point for ‘strongly disagree’. Reverse items (e.g. ‘I don’t worry about making mistakes in the Spanish class’) were also reverse coded.

4.3.2.4.3 Qualitative Measurements: Open-Ended Questions for State Anxiety, Task Difficulty and Mental Effort

Participants’ open-ended responses were coded using qualitative content analysis (Corbin & Strauss’, 2008) to find meaningful categories that emerged from the data. A total of 610 responses were generated for the whole data sample with a number of 12 answers (for each open-ended question) per participant. Responses were coded by the main researcher in three sections in the same logical order as participants completed the questionnaire (questions: 1a, 1b, 2a, 2b, 3a, 3b, 4a, 4b).

Coding for content comprised several stages: accounting for themes specific for each construct separately. Sections 1 and 2 were related to ‘general emotions’ and ‘state anxiety’ (respectively); section 3 examined ‘task difficulty’, and section 4 examined ‘mental effort’. Importantly, the state anxiety, the task difficulty and the mental effort open-ended questions were phrased, coded and analyzed in relation with L2 performance.
Participants were asked to reflect on how their language anxiety, and the factors they perceived to be most difficult (task difficulty) affected their task performance. Likewise, participants’ responses regarding mental effort were coded according to the main variables of interest, namely, a) perceived state anxiety, b) cognition and processing demands, and c) L2 task performance.

Data was coded separately for the cognitive simple and the cognitive complex task for each participant. During the process, a first stage consisted on closely examining responses in detail to come up with emerging themes; a second stage was devoted to further refining the coding frame by adding new categories (not previously identified) to better account for the data. A second coder, an SLA researcher specialized in language anxiety and affective factors, analyzed 10% of the data and came up with her own coding frame. For inter-rater reliability purposes, the second coder was not shown the main researcher’s coding system to avoid biased interpretation of data. After a close examination of the two researchers’ coding frames, data was analyzed according to subordinate, ordinate, and superordinate themes.

In this first step of the coding, subthemes, themes, and superordinate themes were generated to represent participants’ comments. The superordinate themes were color-coded for each answer. Essentially, ‘emotional factors associated to language anxiety’ in the participants’ responses were coded in red, ‘the relationship between language anxiety and cognitive processes’ was coded in green, ‘task-internal features’ in orange, ‘task linguistic demands’ in purple, ‘the relationship between anxiety and L2 task performance factors’ in blue, and task ‘difficulty factors’ were coded in brown.
From more specific to more general, first, the subordinate themes were coded according to these major categories emerging from the extracts. For instance, when the participant mentioned “he felt frustration for not being able to remember vocabulary”, ‘frustration’ and ‘remembering’ were colored in red and green, respectively, as frustration is the specific emotion triggered by ‘not remembering’, a cognitive process. Secondly, for the ordinate themes, ideas were further reduced into the more specific themes, so specifically “frustration” was coded as a “negative emotion associated to anxiety”.

Interestingly, one of the most prominent themes that emerged in this stage and was highly emphasized by the two coders was the relationship between anxiety and cognition, named ‘relationship anxiety-cognition’ and added to the coding framework as it accounted for a large percentage of the data. Emotions and anxiety were revealed in the data as both the cause and the consequence of various cognitive processes, and in turn, were perceived to greatly affect learners’ task performance in Spanish as an L2. As mentioned above, many language anxiety researchers have raised the issue of explaining the directionality of the relationship between language anxiety and L2 performance/achievement to answer the question of whether language anxiety is the cause or the consequence of poor performance. Thus, to account for this relationship, the researcher coded for each time that a) ‘anxiety’ (or other negative emotions) and ‘cognitive processes’ were mentioned together; b) ‘anxiety’ and ‘task performance’ were mentioned together; or c) the three-way interface among anxiety, cognition, and performance were mentioned together.

For the superordinate themes, data was subsequently categorized into the following general and broad content categories: ‘emotions’, ‘relationship Language Anxiety — Cognition — L2 performance’, ‘relationship Language Anxiety — L2 performance’, ‘task-
internal features’, ‘task-linguistic demands’, ‘task difficulty’, ‘mental effort’, and ‘other factors’. ‘Task-internal features’ included comments related to task design features, such as, the storyline, number of characters involved, clarity of pictures, topic, setting of the narrative, length of task; ‘task linguistic demands’ was related to the language that was required to perform the task, specifically grammar and vocabulary, (described by Skehan, 1998, as ‘code complexity’); the ‘emotions’ category comprised all affective factors, emotions and feelings — positive and negative— (expressed either by adjectives or verb forms) identified by participants during their performance.

It is important to clarify that the current dissertation will only report the results and findings pertaining to language anxiety as the affective variable of interest. However, for coding purposes and to investigate the interface Language Anxiety — Cognition — L2 Performance, the two coders created the ‘emotions’ category as an overarching theme. This category accounted for all the positive and negative emotions (other than language anxiety) that were mentioned in participants’ responses (expressed either by adjectives or verb forms). Subcategories included in this superordinate category were: ‘negative emotions’, and ‘positive emotions’. All positive and negative emotions associated with anxiety mentioned in each comment were first counted, and second, listed to keep track of what particular emotions were more recurrent. Specifically, the researcher coded for the number of times that anxiety-related words were mentioned (e.g. frustrated, feeling calmed, worried, feeling anxious, overwhelmed, self-conscious, confident).

For the ‘relationship Language Anxiety - Cognition - L2 performance’, ‘cognitive processes’ such as ‘memory’, ‘thought process’, ‘thought and idea organization’, ‘focus’, etc. were coded in connection with how they were affected by anxiety or how they affect
language anxiety. Often, these subcategories accounted for comments associated with: not remembering the necessary vocabulary or verb tenses, finding it more difficult to organize the ideas to narrate the story, slow thought process or freezing up during task performance due to anxiety, getting side-tracked or distracted, pausing between ideas.

The ‘relationship between anxiety and L2 performance’ was a recurrent theme where participants commented on how task performance was perceived to be affected overall and mostly by language anxiety. This superordinate category included the following subcategories: main affected areas L2 performance, overall good performance, and overall poor performance. The areas mentioned more often by participants were fluency, and verb accuracy. Another theme that was established as a category was ‘task difficulty’ where participants described what aspects of the task were perceived as more complicated and difficult. Often, there was overlapping between the information provided in these comments and the perceived state anxiety.

Overall, new subcategories were generated and used in the coding framework depending on the frequency of occurrence in the data. It is important to note that data was coded in an exhaustive and mutually exclusive way until all the content was represented accurately in the coding frame. These new categories did not fit any of the main existing superordinate categories, and thus, were coded as ‘other factors’. In most cases, these new subcategories referred to other task factors that contributed to triggering state anxiety during task performance and impacted (positively or negatively) L2 performance. The majority of these factors were related to task conditions, and task familiarity issues, and were coded under the following subcategories: a) lack of preparation, b) specific-situation anxiety or speaking anxiety c) time pressure, d) task familiarity (task novelty) e) getting
accustomed to the task as task progressed, f) no note-taking prior to performance; g) distracting environment conditions, and h) task interruption half-way through.

With regards the scoring procedure, if the category was present in the participants’ comment, then it was coded with the number 1, and when it was absent, it remained blank. After exhausting the coding process for each category, numbers were added to calculate how many times that theme was represented for the whole sample data. In those cases where comments were shown contradictory in different answers for the same participant, that category was coded as ‘neutral’.

The next table presents the main categories and some of the most recurrent subcategories with examples that illustrate each category:
Table 3: Superordinate and ordinate categories for the qualitative content analysis with examples extracted from the data.

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotions⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Frustration</td>
<td></td>
<td>&quot;It makes me nervous that I don’t know what words to use, and makes me frustrated when being nervous makes me forget what I was going to say&quot;. (simple task, P. 16)</td>
</tr>
<tr>
<td></td>
<td>Self-consciousness</td>
<td></td>
<td>“This task made me feel a little anxious, and self-conscious, because I’ve studied Spanish for a while, but I still make more mistakes than I probably should.” (complex task, P. 12)</td>
</tr>
<tr>
<td></td>
<td>Lack of confidence</td>
<td></td>
<td>“I think it was rather crushing in confidence that I have this very simple cartoon and yet couldn’t explain it in a story. Having worked in a Spanish speaking environment and spent a month in Panama I feel I have very utilitarian and understandable Spanish, but it was sad to see I couldn’t really expand on that or even really give anything theatrical effect to the story.” (simple task, P. 2)</td>
</tr>
<tr>
<td></td>
<td>Comparison to native speaker skills</td>
<td></td>
<td>“Even though this wasn’t graded, I would have preferred to be a proficient Spanish speaker and so I feel some anxiety and insecurity when I am trying to describe something and don’t have the words to describe something.” (complex task, P. 8)</td>
</tr>
</tbody>
</table>

⁶ Emotions associated with anxiety
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worry</td>
<td></td>
<td></td>
<td>“I think that having more confidence made it easier to communicate smoothly and tell the story, but when I started worrying about my performance it undermined my attempt to communicate.” (simple task, P.2) “I was anxious to forget certain words and was nervous that my Spanish wouldn’t be coherent.” (complex task, P. 11)</td>
</tr>
<tr>
<td>Embarrassment</td>
<td></td>
<td></td>
<td>“I did feel slightly embarrassed that I was having such a hard time describing a simple plot.” (simple task, P. 5)</td>
</tr>
<tr>
<td>Somatic anxiety</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stumbling</td>
<td></td>
<td></td>
<td>“I found it very frustrating that I couldn’t remember all of the vocab that I wanted to use to describe the situation. This frustration served to trip me up and make me more anxious, which made me stumble even more.” (complex task, P. 30)</td>
</tr>
<tr>
<td>Fear of negative evaluation</td>
<td></td>
<td></td>
<td>“I was also nervous speaking out loud due to fear of judgment that might arise from those around me.” (simple task, P. 3)</td>
</tr>
<tr>
<td>Awareness of poor performance</td>
<td></td>
<td></td>
<td>“I didn’t do as well as I think I could’ve.” (complex task, P. 21)</td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td>Good</td>
<td>“The task felt good: the cartoon was much simpler and had actions that I could describe/say in Spanish.” (simple task, P. 19)</td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confident</td>
<td></td>
<td></td>
<td>“The storyline was easy to follow and made me feel <strong>confident</strong>. I had an easier time with the vocabulary.” (simple task, P. 4)</td>
</tr>
<tr>
<td>Fun</td>
<td></td>
<td></td>
<td>“It was <strong>fun</strong> trying to figure out the story with the man and the dog” (complex task, P. 22)</td>
</tr>
<tr>
<td>Engaging</td>
<td></td>
<td></td>
<td>“The story telling task was easy because the storytelling was more <strong>engaging</strong> and provided more information.” (complex task, P. 14)</td>
</tr>
</tbody>
</table>

**Relationship Anxiety-Cognition- L2 Performance**

1. Anxiety affecting cognition (cause)

<table>
<thead>
<tr>
<th>Thought and idea organization</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“The more frustrated/anxious I was, the more difficult it was to <strong>organize my thoughts</strong> and communicate effectively” (simple task, P. 6)</td>
</tr>
<tr>
<td>Focus</td>
<td>“That probably worsened my performance because anxiety made me more self-conscious which made me less <strong>focused</strong> on the task, which also led me to perform worse because of this cycle of anxiety.” (complex task, P 2) “It made it more difficult to <strong>focus</strong> and <strong>concentrate</strong>.” (simple task, P. 9)</td>
</tr>
<tr>
<td>Thought process</td>
<td>“<strong>Being stressed made me freeze up</strong> as I was speaking whenever I didn’t know a word, I <strong>lost my train of thought</strong> and became frustrated.” (simple task, P. 17)</td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td></td>
<td></td>
<td>“I spoke slowly and worried I wasn’t doing well enough. I felt nervous which made it harder to remember things. It was only when I really relaxed did the process become fun, even if I wasn’t able to describe the scene as well as I would have been able to in English.” (complex task, P. 5)</td>
</tr>
<tr>
<td>2. Cognition affecting anxiety (consequence)</td>
<td></td>
<td></td>
<td>“I think that using words I know for a fact I have learned for class brought anxiety on because I knew it was in my head somewhere and I just could not remember them at all.” (simple task, P. 10)</td>
</tr>
</tbody>
</table>

Relationship Anxiety-L2 Performance

<table>
<thead>
<tr>
<th>Areas affected in output modification</th>
<th>L2 fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“My story was longer and discussed the actions of most of the characters in the pictures and their interactions with each other.” (complex task p. 34). “It made my performance bad because I kept pausing and trying to figure out what to say but ended up not saying anything or saying things in English.” (complex task p. 33). “The nervousness made me flustered and there were huge gaps and “ummm”s in my speaking.”’” I spoke in a stilted manner”. (simple task, p. 16)</td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 Grammar</td>
<td></td>
<td></td>
<td>“My <strong>verb tenses</strong> were off. It was difficult to figure out which tense would be correct.” (complex task, P.17.). “For this task, thinking of the terms “walking a dog”, running “away from”, and “swerving” took a long time, and caused me to go slower in describing the sequence of events. I also wanted to be sure to use correct male/female concordancia, which provoked more anxiety.” (complex task, P. 27). “Conjugations was a struggle for me because I went back and forth between imperfect and present” (complex task, P. 28)</td>
</tr>
<tr>
<td>Creativity</td>
<td></td>
<td></td>
<td>“It was <strong>short and extremely uncreative.</strong> Because I was so anxious about grammar + vocab I basically just described what was going on instead of being creative.” (simple task, p.18). “Vocabulary made me less anxious because there were more details, allowing for more creativity.” (complex task, P. 17)</td>
</tr>
<tr>
<td>Facilitative anxiety</td>
<td></td>
<td></td>
<td>“I think that <strong>helped my performance,</strong> because there was more that I could draw from to talk about. By making it more interesting for me, it also led me to focus more constructively.” (complex task, P. 3)</td>
</tr>
<tr>
<td>Debilitative anxiety</td>
<td></td>
<td></td>
<td>“My anxiousness definitely had a <strong>detrimental effect on my performance.</strong> Many of my pauses were a result of me trying to find the right words.” (complex task P. 25)</td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task-internal features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manipulation of +/-</td>
<td>Few elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Characters</td>
<td></td>
<td></td>
<td>&quot;This task was not very anxiety-provoking because it was a story in which there were only a few subjects to be spoken about (the man, clock, window, bed) so a large vocabulary was not necessary.&quot; (simple task, P.28). &quot;The multiple characters made me anxious because I was unclear as to how to differentiate between them.&quot; (complex task, P. 15)</td>
</tr>
<tr>
<td>Storyline</td>
<td></td>
<td></td>
<td>&quot;The storytelling was pretty straightforward. I named the main character &quot;Juan&quot; and talked about him getting ready for bed and then getting up in the morning.&quot; (simple task, P.1). &quot;I felt more overwhelmed in this story because the photos were more complex. It was difficult to find the storyline because there was much more action and background in each photo.&quot; (complex task, P. 13)</td>
</tr>
<tr>
<td>Clarity of pictures</td>
<td></td>
<td></td>
<td>&quot;The images were a bit more complicated and were harder to discern.&quot; (complex task, p.15). &quot;The simplicity of the situation made it easier, but the obscure items in the cartoon made it hard to properly tell the story.&quot; (simple task, P. 10).</td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
<td>&quot;The length of the story made me a little stress as I was trying to prepare but allowed me to provide more detail, decreasing my anxiety. (complex task, P. 15)</td>
</tr>
<tr>
<td>Superordinate Categories</td>
<td>Ordinate Categories</td>
<td>Subordinate Categories</td>
<td>Example</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------</td>
<td>------------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Task-linguistic demands/ Code complexity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>“Not knowing <strong>enough vocab</strong> made the task much harder, because I had to circumlocute the words that I didn’t know, which interfered with telling the story. It also slowed me down and made the task more stressful. (simple task, P.4). “I understood the plot of the story, but since I didn’t know all of the vocabulary (alarm clock; hit; curtains.) so I got hung up. (simple task, P. 2)”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>“The most anxiety-provoking aspect was the vocabulary and <strong>grammar.</strong>” (complex task, P. 13). “The <strong>grammar</strong> and vocab are most difficult and give me the most anxiety provoking because <strong>I can never think of the conjugations quickly</strong>” (complex task, P. 14) “I believe using basic grammar helped make me less anxious because I felt comfortable using these tenses.” (simple task, P. 15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task difficulty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task-internal features/ Conceptual task demands</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+/- Few Elements (+/- Characters)</td>
<td>“The topic was easier. Before it was a man and dog being arrested for running into a couple having coffee. This was <strong>simply about a man</strong> (I character) going to bed and waking up.” (simple task, P. 32).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storyline</td>
<td>“I think that the task was also harder because of <strong>increase in activity in these pictures.</strong>” (complex task, P. 3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td></td>
<td></td>
<td>“The task was easier because the storyline was simple, and it was short.” (complex task, P. 17)</td>
</tr>
<tr>
<td>Simultaneity of cognitive processes &amp; Limited Attentional Resource pool</td>
<td></td>
<td></td>
<td>“To think of a storyline requires creativity, to speak in a foreign language requires good understanding and recollection of vocabulary, to speak these things in Spanish is difficult. This uses two sides of the brain.” (complex task, P. 8)</td>
</tr>
<tr>
<td>Task-linguistic demands/Code complexity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td>“I struggled most with the vocabulary—if I was familiar with the vocab I think the task would have been much easier and more pleasant.” (simple task, P. 1)</td>
</tr>
<tr>
<td>Grammar</td>
<td></td>
<td></td>
<td>“It was difficult to relay my thoughts into a grammatically correct Spanish story.” (complex task, P. 17)</td>
</tr>
<tr>
<td>Other task factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task mode</td>
<td></td>
<td></td>
<td>“It is hard to use other verb tenses when speaking not writing though.” (simple task, P. 35). It is much easier to write it than speak it.” (complex task, P. 17)</td>
</tr>
<tr>
<td>Task repetition</td>
<td></td>
<td></td>
<td>“I think I did better telling the story on this task, because it was simpler and because I had done it before.” (complex task, P. 3)</td>
</tr>
<tr>
<td>Superordinate Categories</td>
<td>Ordinate Categories</td>
<td>Subordinate Categories</td>
<td>Example</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Lack of planning</td>
<td></td>
<td></td>
<td>“I need some <em>preparation</em> before I tell a story. My learning process requires me to write down the information. It’s difficult when it has to be very spontaneous.” (simple task, P. 35).</td>
</tr>
<tr>
<td>Time pressure</td>
<td></td>
<td></td>
<td>“The story telling task was slightly easy because the instructions were simple and the story was simple. However, trying to decide how to tell the story in 30 seconds was difficult because I <em>felt rushed</em>.” (simple task, P. 35)</td>
</tr>
</tbody>
</table>

**Mental effort**

<table>
<thead>
<tr>
<th>Task-linguistic demands/Code complexity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>“The task required more mental effort because I was trying as hard as I could to remember <em>basic vocabulary</em> and I simply could not. Even as I skipped certain details and moved on, I was still thinking back to the prior images trying to think of how I could have forgotten such simple <em>vocabulary</em>”. (simple task, P. 17)</td>
</tr>
<tr>
<td>Grammar</td>
<td>“Trying to make a nice story without making too many mistakes required more mental effort.” (complex task, P. 10)</td>
</tr>
</tbody>
</table>
**Table 3. (cont’d)**

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneity of cognitive processes &amp; Limited attentional resource pool</td>
<td>“The pictures, of course, helped with guiding it. Not being able to take notes beforehand made it more difficult. Thus, I was trying to come up with words to say, <strong>conjugate the verbs</strong>, and also trying to think of my next sentence and where to go with the story, all at the same time.” (complex task, P. 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“The reason why I did not put full mental effort in is because at a certain point I realized I just was not going to be able to explain the story since I knew so few relevant words, and I was not going to experience any consequences for stopping or acknowledging I failed. It required a lot of effort though because there were so many elements to consider, such as grammar, creativity, and basic vocab.”</td>
</tr>
</tbody>
</table>

**Other task factors**

<table>
<thead>
<tr>
<th>Speaking anxiety/ task mode</th>
<th>“I felt as though others were listening to me speak and that made me feel more <strong>stressed</strong>. “Makes me anxious to be speaking in a room in front of other people.” (P. 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task repetition task novelty)</td>
<td>“<strong>I feel unprepared to speak on something I haven’t seen</strong>. I do not have an adequate amount of knowledge about Spanish to speak on a timeline of events.” (complex task, P. 26)</td>
</tr>
<tr>
<td>Superordinate Categories</td>
<td>Ordinate Categories</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Task type</td>
<td></td>
</tr>
<tr>
<td>Lack of planning</td>
<td></td>
</tr>
<tr>
<td>Time pressure</td>
<td></td>
</tr>
<tr>
<td>No note-taking</td>
<td></td>
</tr>
<tr>
<td>Task familiarity as task progressed</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. (cont’d)

<table>
<thead>
<tr>
<th>Superordinate Categories</th>
<th>Ordinate Categories</th>
<th>Subordinate Categories</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distracting environment conditions</td>
<td></td>
<td></td>
<td>“From the beginning of the task, I felt slightly uncomfortable formulating the proper Spanish while being recorded”. (simple task P.8). “The pressure of being recorded and not having the vocabulary available made it difficult.” (complex task, P. 16)</td>
</tr>
<tr>
<td>Task interruption</td>
<td></td>
<td></td>
<td>“Initially I felt very confident, but lost concentration somewhat after pausing in the middle.” (complex task, P. 4)</td>
</tr>
</tbody>
</table>
L2 proficiency and production have been defined as a multifaceted construct with a dynamic interrelationship between its main components (Skehan, 2009), and captured by the notions of complexity, accuracy and fluency—referred to as the triad CAF (Ellis, 2003; 2008; Ellis & Barkhuizen, 2005; Housen, Kuiken & Vedder, 2012). In addition, successful performance in task-based environments has been associated with the more advanced language (complexity), “correctness” or error avoidance (accuracy), and “the ability to produce speech at an appropriate rate and without interruptions” (Skehan, 2009).

For comparability reasons with previous TBLT literature (e.g., Housen, Juiken & Vedder; 2012; Kuiken & Vedder, 2008; Kormos, 2011; Michel, Kuiken & Vedder, 2012; Robinson, 2001b, 2007; 2011; Sasayama, 2016; Tavaloki & Foster, 2008) L2 task performance in the current study was examined by analyzing the learner’s complexity, accuracy and fluency linguistic dimensions in the simple and the complex condition. Audio files—recorded in QuickTime—were transcribed for each participant and each task complexity condition to be subsequently analyzed for CAF indices.

Following Norris and Ortega (2009), Analysis of Speech unit (AS-unit) was selected being one of the most widely used and reliable units of analyses for oral data in SLA research. Specifically, Foster et al. (2000) defined the AS-unit as “a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (Foster et al., 2000, p. 365). The AS-unit is the equivalent to the so-commonly employed T-unit (Hunt, 1970) more widely applied in measuring CAF in L2 written data (although, it has been used in both the oral and written
A large number of performance measures were used to assess oral production in the two task conditions (for a summary, see Table 4).

**Table 4**: Overview of the CAF measures for L2 task performance.

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complexity</strong></td>
<td>syntactic complexity</td>
<td><em>mean length of AS-unit</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>mean length of clause</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>number of clauses per AS-unit</em></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>accuracy</td>
<td><em>weighted clause ratio</em></td>
</tr>
<tr>
<td></td>
<td>&amp; comprehensibility</td>
<td></td>
</tr>
<tr>
<td><strong>Fluency</strong></td>
<td>productivity</td>
<td><em>total number of words</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(excluding repetitions, repairs, reformulations, and false-starts)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>total number of syllables</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(excluding repetitions, repairs, reformulations and false-starts)</td>
</tr>
<tr>
<td></td>
<td>pruned speech rate</td>
<td><em>number of words per minute</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>number of syllables per second</em></td>
</tr>
</tbody>
</table>

### 4.3.2.4.4.1 Fluency

Fluency has been most commonly viewed as “learner’s control over their linguistic L2 knowledge, as reflected in the speed and ease in which they access relevant L2 information to communicate meanings in real time” (Wolfe-Quintero et al. 1998: 4).

As with the CAF triad, fluency in the L2 has posed difficulties in its operationalization as a theoretical construct due to the vast number of existing definitions and current measures that exist today. The many different ways in which fluency has been
conceptualized to date in SLA research have to do with its multidisciplinary nature (Segalowitz, 2010; Skehan, 2009; García-Amaya, 2012).

To be consistent with previous task complexity literature, the current study employed *utterance fluency* as the primary measure to operationalize oral fluency. Utterance fluency is concerned with the physical production and measurable characteristics of speech as a product (often being called temporal phenomena or speed fluency) (Baker-Smemoe et al., 2014; Skehan, 2009). Additionally, the concept of utterance fluency has been further divided into three categories that account for different aspects of speech: *speed fluency* (the rate at which the speech is delivered), *breakdown* (hesitations and pauses) and *repair fluency* (frequency of corrections, self-repairs, false starts and repetitions) (Tavaloki & Skehan, 2005).

To be able to establish comparisons with previous task complexity literature (e.g., Robinson, 2001b, 2007; Sasayama, 2016; Skehan, 2003; Tavaloki & Skehan, 2005), the present study selected several quantitative measures of productivity (amount of language produced) (Norris & Ortega, 2009) and speech rate (Robinson, 2001b, 2007; Sasayama, 2016; Skehan, 2003; Tavaloki & Skehan, 2005).

For productivity, two common productivity measures were chosen: total number of words spoken (Freed et al., 2004; Segalowitz & Freed, 2004), and total number of syllables (Temple, 1992). For speech rate, due to the fact that in Spanish words usually present a substantial variation in number of syllables per word, the present study employed both words and syllables as the main unit of analysis. The rationale behind this decision was to achieve a more exact insight of the learners’ potential gains in their oral production.
Specifically, the following measures were used for speech rate: words per minute (e.g., Segalowitz & Freed, 2004) and (pruned) syllables per second (e.g., Drackert, 2016; Sasayama, 2016). Interestingly, the latter has been found to be one of the most popular and sensitive measures in L2 fluency research. It has been commonly used in study abroad (e.g., Cuchiarini et al., 2000, 2002; De Jong et al., 2009; Derwing et al., 2004, 2009; Tavaloki & Skehan, 2005), and task complexity literature (e.g., Drackert, 2016; Robinson, 2007; Sasayama, 2016).

To calculate the ratios of words per minute and syllables per second, the total number of words and syllables, respectively, was divided by the total number of minutes and seconds, respectively, spent on each of the tasks, the time on task. Importantly, to control for time on task within and between participants’ performances, it was decided to count the number of words and syllables produced in the first 2 and 1/2 minutes (2 minutes and 30 seconds) of all task conditions (except for those tasks that lasted less than 2 and half minutes, where the shortest task—between the simple and the complex—was selected as point of reference). For all productivity and rate of speech measures pauses, repetitions and false starts were not included in the calculation.

4.3.2.4.4.2 Accuracy

Applied linguists and SLA scholars have agreed that accuracy is one of the crucial language dimensions predicting L2 development and acquisition (Norris & Ortega, 2003). To date, it has also been proved to be one the most consistent and straightforward constructs from the CAF triad (Housen & Kuiken, 2009; Housen et al., 2012; Palloti, 2009). Accuracy (or “correctness”) has been referred to “the degree to which L2 learner’s performance and the L2 system underlying this performance deviates from a particular norm (i.e. usually the

Essentially, measures of accuracy have been mainly categorized into two broad types: global measures and local measures of accuracy. Whereas global measures (error-free t-units, error-free clauses, number of errors per 100 words, length of error-free run, ratio of lexical errors per clause) gauge the overall level of lexical or grammatical accuracy in the learners’ L2 production (Skehan & Foster, 2007); local measures tend to capture the developmental stage of acquisition of certain L2 features (i.e. irregular past tense, the use of a particular verb form, subject-verb agreement) or explore how learners respond to specific pedagogic treatments. This has often been done by examining target-like uses of certain morphemes, analysis of obligatory (and non-obligatory) contexts of the use for particular grammatical features and forms etc. (Foster & Wigglesworth’s, 2016). Although these two distinct categories of measures have been widely employed and serve different purposes, lately SLA researchers have pointed out some weaknesses in the operationalization of the construct.

First, according to Foster and Wigglesworth (2016), a common challenge encountered by researchers measuring accuracy is the difficulty in error quantification when a global measure is used. In this case, calculating mean number of errors per unit of analysis has posed some problems in identifying the number of errors produced and the most appropriate way to correct the error. An illustrative example of this phenomenon was taken from the data:

“*El hombre necesita cierra la ventana.*” (‘the man needs *close the window*) (participant 5, simple task)
It is obvious that there are several correct versions and ways to provide feedback to the errors present in the previous clauses regarding the correct past tense form: (a) “El hombre tuvo que cerrar la ventana” (the man had to close the window), (b) “el hombre cerró la ventana” (the man closed the window), (c) “el hombre tuvo la necesidad de cerrar la ventana” (the man needed to close the window) etc. Not only are there different possibilities to provide feedback to this sentence, but it also seems tricky to determine the exact number of errors. Is it 2 errors, one on necesito (vs necesitó) and another cierra (vs the infinitive cerrar), or should “necesita cierra” simply be considered as one single error?

The most important advantage of this measure is that it takes error gravity into consideration by weighting the error as more or less severe, something most global and local measures fail to address. In general, global accuracy measures usually employ binary distinctions classifying units of analysis as error-free or not error-free (Foster & Wigglesworth, 2016). However, it is obvious that not all errors can be rendered the same importance. The examples below illustrate two clauses that contain a minor error that does not affect communication, and another containing a severe error that does affect communication.

(a) The man go to jail. (1)

(b) The man is jail. (1)

Although examples (a) and (b) contain the same number of errors, one each, giving a score of 1 point in both cases would be problematic since it does not capture the gravity of the error. Evaluating the severity of an error becomes extremely important for language instructors when they provide feedback to learners.
Foster and Wigglesworth (2016) recently developed a new accuracy measure that seems to have addressed (at least partially) the concern of the gravity of the error, the weighted clause ratio (WCR). The current study has employed this accuracy measure to assess participants’ accuracy levels in the simple and complex task. Below, a detailed explanation is given on how units of analysis (clauses) were categorized for this measure, the procedure for how errors were weighted and the scoring system.

Table 5: Clause categorization for weighted clause ratio.

<table>
<thead>
<tr>
<th>Clause Category</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entirely accurate</td>
<td>The clause is accurately constructed.</td>
</tr>
<tr>
<td>Level 1</td>
<td>The clause has only minor errors (e.g. in morphosyntax) that do not compromise meaning.</td>
</tr>
<tr>
<td>Level 2</td>
<td>The clause contains serious errors (e.g. verb tense, word choice or word order) but the meaning is recoverable, though not always obvious</td>
</tr>
<tr>
<td>Level 3</td>
<td>The clause has very serious errors that make the intended meaning far from obvious and only partly recoverable.</td>
</tr>
</tbody>
</table>

Foster and Wigglesworth (2016)

Foster and Wigglesworth’s (2016) weighting error scheme and scoring system were adopted to calculate accuracy indices. AS-units and clauses were used as primary units of analysis, and data was analyzed by two coders for inter-rater reliability for each task. The second coder was a TBLT researcher in the area of task complexity familiar with this particular measure, and coded the data for 10% of data. The two coders participated in a two-hour coding training to ensure agreement in the error weighting and the scoring system for the different levels of error type clauses. Importantly, researchers agreed on how they were going to score clauses according to the specific population, the participants’ levels of proficiency in Spanish, the nature of the task, the complexity of the two task conditions, and the complexity of the linguistic code (i.e. grammar and vocabulary) elicited by the task.

126
To calculate the inter-rater reliability for this measure, all clauses where a rating disagreement occurred were added up, divided by the total number of clauses, and multiplied by 100. A 95.51% was agreed upon between the two coders for the weighted clause ratio.

First, to assess the severity of the error for each participant, all clauses received an error level, and a score assigned depending on the error level according to Foster and Wigglesworth’s scoring system (2016) (see Table 6 for scoring scheme below).

**Table 6: Scoring system and error weighting for weighted clause ratio.**

<table>
<thead>
<tr>
<th>Clause Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entirely accurate</td>
<td>1</td>
</tr>
<tr>
<td>Level 1 error</td>
<td>0.8</td>
</tr>
<tr>
<td>Level 2 error</td>
<td>0.5</td>
</tr>
<tr>
<td>Level 3 error</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Foster and Wigglesworth (2016)

Second, the total number of clauses in each error level was calculated and multiplied by the specific score attributed to that level (1, 0.8, 0.5, 0.1). This yielded the average scores per error level. The ratio (WCR) (ranging from 0 to 1) was obtained dividing the total raw scores (the sum of all error levels altogether) by the total number clauses. Thus, a WCR score of 1 would only be possible if all clauses received 1 (entirely accurate score), unlikely even for advanced L2 learners, who would probably still make minor errors in certain verb conjugations or lexical forms. A ratio ranging between 0.7 to 0.9 meant the majority of clauses for that task received a score of level 1 error (0.8); a WCR ranging between 0.4 and 0.6 were scored with a majority of level 2 error type in most of the clauses;
and finally, a WRC with a score that ranged between 0.1 and 0.3 means most clauses have received the level 3 error, and thus, performance contained severe errors. It is fairly common that scores actually fall in between those ranges, where a mix of error levels is produced; certain clauses being entirely accurate (or almost accurate) while others contain more severe errors.

In the present study, it was decided that words or constructions used in the participant’s native language, English, would be coded as level 3 errors (despite meaning could be recovered by the coders who were bilingual). Table 7 below shows the weighted clause ratio calculated after scoring on the transcript for the simple task in participant 3 (pruned of pauses, repetitions, and false starts).

**Table 7: Example of categorization and scoring.**

<table>
<thead>
<tr>
<th>Clause</th>
<th>Error level</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1   En la primera fotograf, el hombre es muy cansado</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>2   y es tiempo para noche</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>3   el hombre necesito su clock o tiempo para la mañana</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>4   si, en el segundo fotó el hombre quiere cosa con a ventana</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>5   y también el hombre ve estrellas y la luna</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>6   y en el, finalmente el hombre es muy cansado también</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>7   y también es noche</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>8   y yo, el hombre es <em>jesus, asleep</em></td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>9   Ahora es la mañana</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>10  lo tiempo para el hombre empezar su día</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>11  el de clock am es gritando</td>
<td>3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Accuracy rating: 5.1: 11= 0.46

**4.3.2.4.3 Complexity**

Linguistic complexity (as it occurs with fluency) is a multifaceted construct concerned with various linguistic qualities of the language. In this sense, a fundamental
distinction needs to be made when this construct is addressed. Whereas syntactic complexity is viewed as learners’ abilities to use a wide range of sophisticated structures, lexical diversity (or variety) considers the ability to use a wide range of vocabulary in the L2 (Bulté & Housen, 2012; Ellis & Barkhuizen, 2005; Plonsky & Kim, 2016).

In the current study, the researcher examined syntactic complexity as the main focus, employing different measures with the purpose of accounting for the various dimensions of syntactic complexity. Following Norris and Ortega’s (2009) suggestions’, a combination of a) length-based measures (number of words, syllables, morphemes, or characters per unit of analysis)—a global measure of syntactic complexity-, b) amount of subordination, and c) complexity measures at the phrasal level (mean length of clauses) are recommended for various reasons.

First, these measures capture different but complementary dimensions of the same construct, since the first two types tap structural complexity at a clausal (or inter-sentential level); and the latter specifically looks at sub-clausal complexity. It is important to realize that clause length does not depend on subordination, and so they remain independent syntactic processes since the latter will not affect the former (Norris & Ortega, 2009). Additionally, the use of these measures together has allowed SLA research to distinguish L2 performances at different proficiency levels. In this regard, Norris and Ortega (2009) argued that while length-based measures are useful to distinguish among various proficiency levels, amount of subordination has been found appropriate particularly at lower levels of proficiency (introductory and intermediate); and mean length of clause discriminated among more advanced levels of proficiency. Specifically, the authors explained that lower proficient learners tend to produce more subordination as the main
source of syntactic complexification. On the contrary, at a more advanced developmental stage, L2 learners usually tend to generate more nominalizations (more complex nouns or adjectival phrases), which result in structural changes at the phrasal level. In this regard, a trade-off often occurs where learners complexify language more at the clausal level, and thus, amount of subordination is expected to decrease (Norris & Ortega, 2009).

Therefore, three measures were selected: 1) mean length of AS-units (MLASU) (Hunt, 1970), a measure of global syntactic complexity; 2) clauses per AS-unit (C/A) (Elder & Iwashita, 2005; Skehan & Foster, 2005; Michel et al., 2007), used to measure amount of subordination; and mean length of clause (MLC), primary measure of subclausal complexity (Drackert, 2016; Sasayama, 2016).

For the Mean length of AS-units was calculated by diving the total number of words produced by the total number of AS-units. For amount of subordination, averages of the total number of subordinate clauses per AS-unit; and finally, to calculate clauses per AS-units, clauses were counted and, then, divided by the total number of AS-units.

Importantly, the division of data into AS-units and clauses as main units of analysis allowed ratios as primary measures of syntactic complexity to be calculated. Moreover, following Foster, Tonkyn and Wigglesworth (2000) the AS-unit was considered the most appropriate unit of analysis (instead of the popular T-unit, more commonly employed in written data) given the fragmentary nature of oral data.

An AS-unit, as defined by Foster et al. (2000) is a “single speaker’s utterance consisting of an independent clause or sub-clausal unit, together with any subordinate clause(s) associated with either” (p. 366). In addition, it is essential to understand that, specifically for spoken data, while an independent clause refers to “any phrase clause that
minimally includes a finite verb”, it is fairly common to find independent sub-clauses, phrases that omit conversational elements and still constitute clauses (Foster, 2000: page 369). For clause computation, an example of recovered clause via ellipsis has been coded as follows:

“El hombre y el perro corrió rápido” (the man and the dog run fast) (1 clause, and 1 as-unit).

“Y el hombre y la mujer también.” (and the man and the woman too) (1 clause, 1 AS-Unit) (the verb phrase “running fast” is recovered via ellipsis, and thus, this still constitutes an independent clause).

Following Foster et al. (2000), participants’ utterances were divided as AS-units when they were separated by more than 0.5 seconds or intonational changes. For inter-rater reliability, a second coder analyzed the 10% of the data. A 93.50 for AS-Units and 95 for clauses were the agreement percentages.

4.3.2.5 Data Analyses

Scores from self-assessment measures—state anxiety, classroom anxiety, task difficulty, mental effort and time estimation—were entered into SPSS together with the L2 performance linguistic measures (CAF).

Data analysis was conducted by a combination of descriptive calculations, graphical observations, confidence intervals, inferential analyses and effects sizes. SLA researchers’ recommendations (Norris, 2015; Plonsky, 2013; Tavachnick & Fidell, 2013) on the risks of solely relying on p values and inferential statistics (often leading to problematic interpretation of findings) suggested the important benefits of using various analytical procedures. Inferential analyses and effect sizes—conducted on each of the L2
measures (cognitive load, language anxiety, and CAF linguistic measures)—were interpreted with caution due to the exploratory nature of the study. A One-Way Repeated Measures ANOVA with cognitive task complexity as the within-subject factor (with two levels) was used to confirm if overall statistical differences existed. Lastly, effect sizes employing $d$ values were performed to show the practical application of the results obtained and for the use of future meta-analytical research.

In order to investigate the research questions put forth above, a variety of quantitative and qualitative analytical approaches were conducted on the above measures. First, reliability analyses using Cronbach’s alpha coefficients were calculated on the items for the self-rating language state anxiety measure (at Time 1 and Time 2), the FLCAS (the *Foreign Language Classroom Anxiety Scale*); and to assess participants’ consistency of responses on the self-assessment cognitive load measures—task difficulty, mental effort, and time estimation. Second, scores were inspected employing descriptive statistics and graphical analyses. In particular, data was observed for distributions of normality, examining skewness and kurtosis, and the presence of outliers. Descriptive statistics were explored for means, standard deviations, 95% confidence intervals; and effect sizes were calculated on each independent cognitive complexity and linguistic measure.

Quantitative analyses were performed for the independent complexity measures, the language anxiety measures, and the linguistic outcome measures. For the qualitative analyses, participants’ comments were examined employing content analysis, specifically, grounded theory (Corbin & Strauss, 2008) on the transcribed data from the open-ended questions.
Since one of the main variables of interest was a learner’s individual factor, namely, language anxiety, a within-group design was selected to help reduce errors associated with individual differences. One of the most significant benefits of employing this type of experimental design is that it does not require a large number of participants.

The research design included one independent variable: cognitive task complexity (with two levels: simple vs complex) and several dependent variables: (a) perceived state anxiety (measured at two points in time: Time 1 and Time 2, (b) three cognitive load measures: perceived task difficulty, perceived mental effort, and time estimation; and (c) eight task performance measures for the CAF triad: four measures of oral fluency, one measure for accuracy, and three measures of syntactic complexity.

First, RQ1 validated the actual cognitive demands on the two tasks task (construct validity) and investigated the relationship between cognitive task complexity and cognitive load (Révész et al., 2016; Sasayama, 1026). This was done to ensure that the designed-to-be simple and the-designed-to-be complex tasks were actually imposing higher and lower cognitive task demands for the Spanish L2 learners. Therefore, within-group comparisons were also performed to look at the relationship between the three cognitive load measures (i.e. perceived task difficulty, perceived mental effort, time estimation) and the designed task complexity (as cognitively simple and cognitively complex). Based on the coding frame, qualitative content analyses were performed on learners’ responses to the open-ended questions to provide more detail on their perceptions about what task factors contributed the most to task difficulty and mental effort in relation to the two levels of cognitive task complexity (RQ2a). Inter-rater reliability analyses were conducted on a 10%
of the data, the two coders agreeing on all the major superordinate main ordinate categories and subcategories.

Second, Means, $SDs$, the upper 95% confidence intervals and effect sizes were compared to see if meaningful differences existed between the simple and the complex task on each dependent variable. In addition, within-group differences were examined via One-Way Repeated Measures ANOVA Analysis of Variance (ANOVA) to confirm the previous analyses between the cognitively simple and the cognitively complex task in relation to perceived state anxiety (in Time 1 and Time 2) (RQ2a). Repeated measures ANOVA is the appropriate analysis since all participants were tested under two different conditions (two levels of task complexity). For all cognitive load and linguistic measures, cognitive task complexity was the within-subject factor. For state anxiety, both task complexity and Time were within-subject factors, as participants were assessed in two task conditions and at two points in time (Time 1 & Time 2). Qualitative content analyses on the open-ended questions and inter-rater reliability analyses were performed on the coding frame to interpret learners’ comments on their perceived state anxiety in relation to the two tasks’ cognitive demands (RQ2a). As part of RQ2a, learners’ responses were analyzed to achieve a better understanding on which particular task factors have a greater influence on inducing learners’ state anxiety during task performance.

The foreign language classroom anxiety scores (FLCAS) were employed as a baseline to examine whether a relationship existed between learners’ levels of perceived state anxiety and the cognitive demands of the two tasks. The FLCAS provided scores for the participants’ perceptions on their general classroom predispositions to feel anxiety in the Spanish classroom. Correlational analyses were performed separately on each cognitive
complexity level to investigate the relationship between perceived classroom anxiety scores and perceived state anxiety scores (RQ2b). Pairwise correlations were conducted for Time 1 (half-way through the task) and Time 2 (after the task) on each task to examine if possible minimal fluctuations in the perceived state anxiety occurred. Due to the particular nature of affective factors (constantly varying), arguably, anxiety levels as measured right after task completion could have decreased as compared with anxiety levels occurring right in the middle of the task.

To respond to RQ3, descriptive statistics and effect sizes and inferential analyses (Repeated Measures ANOVAS) were calculated to examine whether language anxiety fluctuate over time when cognitive task complexity was manipulated during oral task performance.

On the three-way interface Language Anxiety—Cognition—L2 Task Performance (RQ4), the relationship between levels of perceived state anxiety and cognitive processes (RQ4a), and the relationship between perceived state anxiety and task performance in L2 Spanish (RQ4b) was examined again by employing qualitative content analysis on participants’ responses to the open-ended questions from the post-task measure, the self-assessed language anxiety questionnaire.

Lastly, descriptive statistics, 95% confidence intervals, effects sizes, and within-group comparisons using significance testing procedures were conducted to explore if meaningful differences existed between the cognitively simple and the cognitively complex task in relation to the L2 performance measures (RQ5).
Table 8: Overview of analytical procedures.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research Question 1</strong></td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td>Inferential analyses: within-group comparisons</td>
</tr>
<tr>
<td></td>
<td>Effect sizes: Cohen’s d (1988)</td>
</tr>
<tr>
<td></td>
<td>Inter-rater reliability analysis for the coding framework for qualitative data analyses</td>
</tr>
<tr>
<td></td>
<td>Qualitative content analysis (grounded theory)</td>
</tr>
<tr>
<td><strong>Research Question 2</strong></td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td>Inferential analyses: within-group comparisons</td>
</tr>
<tr>
<td></td>
<td>Effect sizes: Cohen’s d (1988)</td>
</tr>
<tr>
<td></td>
<td>Inter-rater reliability analysis on the coding framework for qualitative data analyses</td>
</tr>
<tr>
<td></td>
<td>Qualitative Content Analysis (grounded theory)</td>
</tr>
<tr>
<td></td>
<td>Correlational analyses</td>
</tr>
<tr>
<td>Research Question</td>
<td>Analyses</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **Research Question 3** | Descriptive statistics  
Inferential analyses: within-group comparisons  
Effect sizes: Cohen’s d (1988) |
| **Research Question 4** | Qualitative Content Analysis (grounded theory) |
| **Research Question 5** | Descriptive statistics  
Inferential analyses: within-group comparisons.  
Effect sizes: Cohen’s d (1988)  
Inter-rater reliability analysis for accuracy and syntactic complexity measures. |
CHAPTER 5: RESULTS

Results of the data analyses are reported below in the same numerical order of the above research questions. To explore the overall relationship between language anxiety, cognitive processing and task performance in the L2, the following two-way relationships were explored separately.

5.1 Results for the Cognitive Load Measures

The results pertaining RQ1 on the relationship between cognitive task complexity and cognitive load—as measures by task difficulty, mental effort and time estimation are addressed in what follows.

5.1.1 Quantitative Results for the Cognitive Load Measures

5.1.1.1 Task Difficulty

Reliability analyses were performed to examine how consistent participants’ responses were for the three self-assessment measures. For task difficulty, Cronbach’s Alpha was 0.67, showing that it was relatively consistent in distinguishing among participants for the simple and complex task.

For task difficulty, descriptive statistics (see Table 9) were calculated for participants’ self-assessed task difficulty on the cognitively simple and the cognitively complex task. Before exploring frequencies, data was inspected for distributions of normality.

Two missing cases were revealed in the data (N=50). As displayed in Figure 4, participant 39 fell outside the 95% intervals for the two tasks, and thus, the researcher decided to eliminate this case from the data set as it was a bivariate outlier. Participant 12 scored below the lower bound 95% confidence intervals, was a univariate outlier so it was
decided to keep this case in the data set.

Figure 4: Boxplots for task difficulty before outliers were removed based on the two cognitive task complexity levels.

On a scale 1.00 to 6.00, scores ranged between 3.00 and 6.00 for the simple task, showing slightly higher variability than for the complex task, which ranged between 2.00 and 6.00. For the cognitively simple task (see Figure 5), participants’ scores were placed towards the higher end of the distribution, showing data was slightly negatively skewed (-0.03) as compared to the means ($M = 4.38$).
Likewise, for the complex task, a light negative skewness was observed for a mean score of 4.40 (see Figure 6). However, the skewness value (-0.39) was not greater than twice the standard error (0.33), and thus, distributions were considered fairly normal.
Figure 6: Histogram for task difficulty for the cognitively complex task.

Overall, descriptive statistics showed participants perceived the two tasks to be equally difficult, since most of the scores fell in the upper half of the scale. As can be seen in Table 9, means and SDs seemed to indicate no meaningful differences existed between how difficult learners perceived the two tasks nonetheless. The cognitively complex task ($M = 4.40, SD = 1.01$) was only slightly higher than the cognitively simple ($M = 4.38, SD = .92$). However, this difference was not statistically significant. The majority of the participants scored above 4.00, demonstrating that they believed the two tasks were ‘moderately difficult’.
Table 9: Descriptive statistics for task difficulty for the simple and the complex task.

<table>
<thead>
<tr>
<th>Task difficulty</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.38</td>
<td>4.40</td>
</tr>
<tr>
<td>SD</td>
<td>0.92</td>
<td>1.01</td>
</tr>
<tr>
<td>Min</td>
<td>3.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Max</td>
<td>6.00</td>
<td>6.00</td>
</tr>
<tr>
<td>95% Confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval</td>
<td>Lower bound</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>Upper bound</td>
<td>4.64</td>
</tr>
</tbody>
</table>

5.1.1.2 Mental Effort

A different pattern was revealed for mental effort when examining results. Reliability estimates revealed a Cronbach’s Alpha value of 0.67, indicating again a reasonable consistency among participants’ scores between the two tasks. Data was inspected for outliers and participant 42 was removed as a univariate outlier in the complex task. Participants 12, 29, 39 and 50 turned out to be univariate outliers for the simple task. However, it was decided to keep them since they were not extreme values and still represented a significant portion of the data set for the simple task complexity level.

Again, scores were distributed over a 6-score range, with a minimum value of 2.00 and maximum of 6.00 for the simple task; and 3.00 and 6.00 for the complex level of cognitive task complexity. As it occurred with task difficulty, scores were placed slightly towards the upper part of the scale (see Figures 7 & 8). However, when comparing the
numerical values for skewness (-0.44) (simple task), (-0.1) (complex task), these did not exceed twice the standard error (0.33) (simple task), (0.33) (complex task). Thus, skewness was not severely violated, and distributions were considered quite normal.

**Figure 7:** Histogram for mental effort for the cognitively simple task.
For mental effort, mean scores indicated participants perceived a meaningful increase in the amount of brain power used when completing the complex task ($M = 4.76$) as compared to the cognitively simpler task ($M = 4.30$). This time, the differences between the cognitively simple and the cognitively complex tasks were statistically significant. In Table 10, it can be observed that the means for the complex task ($M = 4.76$) is greater than the upper bound 95% CI of the simple task and lies outside of it, providing evidence that differences in mental effort were meaningful.
Table 10: Descriptive statistics for mental effort on the simple and the complex task.

<table>
<thead>
<tr>
<th>Mental Effort</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.30</td>
<td>4.70</td>
</tr>
<tr>
<td>SD</td>
<td>1.11</td>
<td>0.77</td>
</tr>
<tr>
<td>Min</td>
<td>3.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Max</td>
<td>6.00</td>
<td>6.00</td>
</tr>
<tr>
<td>95% Confidence Lower bound</td>
<td>3.98</td>
<td>4.54</td>
</tr>
<tr>
<td>Interval</td>
<td>Upper bound.</td>
<td>4.61</td>
</tr>
</tbody>
</table>

To examine if statistical differences existed between the cognitively simple and the cognitively complex task for task difficulty and mental effort, a One-Way Repeated Measures ANOVA was run separately on each dependent variable.

Results confirmed previous descriptive statistics findings showing a statistical significant difference between the cognitively simple and the cognitively complex task for mental effort ($N = 50$) $F (1, 49) = 12.22, p = 0.001$, partial Eta squared = 0.20, as a construct that plays a role in explaining cognitive task complexity, with a medium effect size of $d = 0.52$ (Cohen, 1988). For task difficulty, significance testing confirmed the previous results where no statistical effect was found ($N = 50$) $F (1, 49) = 0.019, p = 0.89$.

Importantly, an overlapping between the mean scores and the confidence intervals (see Figure 9) suggested a small effect size, and therefore, no meaningful difference was assumed in this case. On the other hand, error bars supported the statistical significant
difference observed in the frequency values between the two tasks’ complexity levels regarding the amount of mental effort exerted. Mean scores for the complex task were revealed higher than the upper bound 95% confidence intervals in the simple storytelling task (see Figure 10 below).

Figure 9: Self-assessed ratings for task difficulty based on the two cognitive complexity levels.
Figure 10: Self-assessed ratings for mental effort based on the two cognitive complexity levels.

5.1.1.3 Time Estimation

For the time estimation measure, a high Cronbach’s Alpha of (0.68) indicated participants were reasonably consistent between their self-ratings for the simple and the complex task. After that, data was explored for missing data and distributions of normality. One participant’s scores were missing. A total of two extreme outliers were identified (see Figure 11). Specifically, cases 1 and 33 were extreme outliers situated very far from the upper bound 95% confidence intervals for the two tasks. For this reason, they were eliminated from the dataset.
Participant 33 was revealed as a bivariate non-severe outlier. Although this case was not an extreme outlier, it was also still situated far from the means for the two tasks and also excluded. After withdrawing outliers, means values, SDs, and CIs demonstrated that the distributions of scores were within the normality range. For this measure, the difference between participants’ judged time and their actual time on task was calculated for the two levels of task complexity. Graphs 12 and 13 displayed the distribution of scores for the two tasks visually.

Figure 11: Time estimation distribution of scores before withdrawing outliers based on the two cognitive complexity levels.
Figure 12: Histogram after outliers were excluded for the cognitively simple task on the time estimation measure (secs).
Figure 13: Histogram after outliers were excluded for the cognitively complex task.

As it can be observed in Table 1, no statistically significant differences were revealed between the simple (\( M = -13.93 \)) and the complex task (\( M = -7.44 \)) for time estimation measure between the two storytelling tasks. The means for both the simple and the complex tasks indicated a negative difference between the estimated time and the actual time on task. However, the mean for the simple task was situated slightly below the means of the complex task.

Results showed that on average, cognitive task complexity did not seem to impact learners’ perception of time spent on the storytelling as compared to the actual time on task in the current study. This trend is also observed in Figure 14 visually.
Table 11: Descriptive statistics for time estimation on the simple and the complex task.

<table>
<thead>
<tr>
<th></th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-13.96</td>
<td>-7.44</td>
</tr>
<tr>
<td>SD</td>
<td>104.06</td>
<td>141.51</td>
</tr>
<tr>
<td>Min</td>
<td>-193</td>
<td>-357</td>
</tr>
<tr>
<td>Max</td>
<td>227</td>
<td>340</td>
</tr>
<tr>
<td>95% Confidence</td>
<td>-44.49</td>
<td>-48.99</td>
</tr>
<tr>
<td>Interval</td>
<td>16.61</td>
<td>34.10</td>
</tr>
</tbody>
</table>
A One-way Repeated- Measures ANOVA confirmed the results revealed in the descriptive statistics as far as the effectiveness of the time estimation measure to capture the actual cognitive demands of the two tasks. A $p$ value greater than 0.05 indicated no statistical significant differences existed between the sample and the complex task ($N = 47$) $F (1,46) = 0.134, p = 0.71$, and thus, the null hypothesis was not rejected in this case.
5.1.2 Testing the Effect of Task Order on the Cognitive Load Measures

The possible effect of task order was examined to determine whether carryover effects occurred for the simple and the complex stories on the cognitive load scores reported by participants. All 51 participants were divided into two conditions. Condition 1 corresponded to those participants that completed the complex task in the second place (simple first – complex second), and condition 2 was the other half of the participants, who performed the complex task in the first place followed the cognitively simple narrative (complex first -- simple second).

First, data was inspected for outliers and distributions of normality on the two conditions. For task difficulty, no outlier was revealed. Only slight negative skewness was found for the two task orders on both the cognitively simple and the cognitively complex task.

Descriptive statistics and inferential analyses indicated that whereas the order of tasks had a meaningful effect for task difficulty, it did not seem to have the same effect for mental effort, as perceived by participants.

Table 12 shows that when the complex storytelling was completed right after the simple one (condition 1), participants perceived the simple task to be statistically more difficulty ($M = 4.92$) than when the order was reversed on condition 2 ($M = 3.95$). For the complex task, the opposite pattern was observed, but differences between groups were not statistical, underscoring, hence, the benefits only occurring on the simple narrative.
Table 12: Descriptive statistics for the order of tasks on task difficulty for the two levels of task complexity.

<table>
<thead>
<tr>
<th>Order of tasks</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S - C</td>
<td>C - S</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>Mean</td>
<td>4.92</td>
<td>3.95</td>
</tr>
<tr>
<td>SD</td>
<td>0.89</td>
<td>0.95</td>
</tr>
<tr>
<td>Min</td>
<td>3.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Max</td>
<td>6.00</td>
<td>5.00</td>
</tr>
<tr>
<td>95% Confidence</td>
<td>4.56</td>
<td>3.55</td>
</tr>
<tr>
<td>Interval</td>
<td>5.28</td>
<td>4.36</td>
</tr>
</tbody>
</table>

The same results were confirmed by the Mixed-design Repeated Measures ANOVA, where an interaction effect was revealed between the between group factor—the order of tasks—and participants’ task difficulty scores ($N = 50$) $F (1,48) = 15.850$, $p = 0.000$, partial Eta squared = 0.24

For mental effort, no bivariate outliers were observed for condition 1 or 2 on the two levels of task complexity. Distributions of scores were slightly skewed but situated within the range of normality. Opposite to the pattern revealed for task difficulty, descriptive statistics (see Table 13) seemed to demonstrate that the order in which the two tasks were performed did not affect participants’ perception of how much brain power they exerted to narrate the stories. On a 1 to 6 scale, non-statistically significant differences were observed between the two task orders.
Table 13: Descriptive statistics for the order of tasks on mental effort for the two levels of task complexity.

<table>
<thead>
<tr>
<th>Order of tasks</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S - C</td>
<td>C – S</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>Mean</td>
<td>4.50</td>
<td>4.25</td>
</tr>
<tr>
<td>SD</td>
<td>1.02</td>
<td>1.15</td>
</tr>
<tr>
<td>Min</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Max</td>
<td>6.00</td>
<td>6.00</td>
</tr>
<tr>
<td>95% Confidence</td>
<td>4.08</td>
<td>3.76</td>
</tr>
<tr>
<td>Interval</td>
<td>4.91</td>
<td>4.73</td>
</tr>
</tbody>
</table>

A lack of effect was equally confirmed by inferential analytical procedures on participants’ scores for self-assessed mental effort (N = 50) $F(1,48) = 0.733, p = 0.396$, partial Eta squared = 0.015

5.1.3 Qualitative Results for Task Difficulty and Mental Effort

To respond to RQ1 on the relationship between cognitive task complexity and cognitive load in L2 Spanish—as measured by subjective measures, for task difficulty and mental effort—as perceived by learners—participants’ comments in the open-ended questions were examined closely according to theme categories extracted for the qualitative coding frame.

Percentages were quantified for each category to calculate the how much each category was represented in the data and contributed to explain the constructs.
Participants’ introspective comments revealed insightful findings that led to a fuller understanding of the relationship between cognitive task complexity, task difficulty and mental effort. Although some participants provided a short comment of one or two sentences, some of them wrote a longer and more detailed explanation of why and how the difficulty experienced affected their performance in the two storytelling tasks. In what follows, the results found in relation to the qualitative data will be examined further.

Aligned with the categories extracted in the coding frame, results in the qualitative data mostly revealed information regarding a) factors that contributed to task difficulty and mental effort; b) whether the tasks were considered easy or difficulty overall; and whether they required much mental effort or not; c) whether task performance was affected by it, and finally, d) how and what aspects of performance were affected in relation to task difficulty.

5.1.3.1 Task Difficulty

With regards to the factors that contributed to task difficulty, participants mentioned a variety of factors in order of importance as how much they were present in the data. The most prominent factors in order of importance were: task-linguistic factors in 78% of all instances, followed by task-internal features represented in 56% of participants’ comments, and other factors related to task conditions and situational factors, that represented 51% of the data.

The category ‘task-linguistic demands’, referred by Skehan (1996) as ‘code complexity’ was the category mentioned by the majority of participants in relation with the challenges faced regarding the difficulty of the vocabulary and grammar elicited in both tasks. Among these two, the impossibility to come up, think of or use the “vocabulary”
required was counted in 74% of the participants’ comments, demonstrating to be by far the most influential factor that contributed to task difficulty for both the simple and the complex task. A representative example of vocabulary as one of the main contributors to task difficulty can be seen below:

“The story-telling task was difficult because I did not have the vocab to express my ideas coherently. I also did not have time to fully organize my thoughts before telling the story.” (participant 36, complex task).

However, let us consider the following example where the participant expressed vocabulary as being less of a problem in the simple task and aiding learner’s performance:

“There was less in the photos. In the first task, there was a lot going on everywhere. In this task, it was just a man going to bed and getting angry at his clock. The second story involves a lot less required vocabulary, so it feels like less of a failure, because I was able to cover more of the story, proportionally speaking.” (participant 25, simple task).

After vocabulary, grammar, and particularly, the accurate use of conjugations and tenses was reported by participants as the second most prominent ‘task-linguistic’ factor that led to the tasks being considered as difficult. These two factors were reported in 21% of the responses for both the simple and the complex task.

In relation to the category ‘task-internal features’ (referred by Robinson, 2005, as cognitive resources), as expected, results demonstrated that +/- few elements—the main operationalization of task complexity employed in the current study—was the second most influential factor, after linguistics demands, represented in 56% of the participants’ responses. In other words, increasing the number of elements and characters in the narrative story (which represents ‘conceptual complexity’) led to a statistical significant increase in
the task difficulty measure between the simple and the complex task.

The +/− few elements and the storyline/amount of action (represented in the coding frame as separate categories) were conflated (and computed together) for analytical purposes, as data revealed that in the majority of the cases participants interpreted and referred to these two interchangeably as the same task feature. Thus, the ‘highest-frequency’ factor was +/− elements, sometimes expressed “the higher or lower number of characters present in the story”, others reported as “more or less elements in the story”, and others as reported by participants as influencing ‘task difficulty’.

The following excerpt reflects the +/− element operationalization of task complexity leading to an increase in task difficulty as perceived by the learners:

“Telling the story was difficult because I wasn't able to comprehend what was happening in the pictures fast enough to tell the story. The multiple things going on in each picture made it hard to focus on the central story. As I'm typing this I'm starting to actually understand what was happening in the pictures, which shows that I wasn't able to comprehend it fast enough.” (participant 48, complex task).

Critically, although for many participants fewer elements present in the story induced less processing, and hence, led them to perceive the simple task as less complicated, for many having less characters and less action going on was a source of difficulty as it to be less creative. In this respect, the complex task allowed them to be more creative as more characters and more elements allowed them to focus on other alternatives whenever they got stuck on certain unknown vocabulary.

“When I lacked vocabulary I still stumbled but I think it was less than the first time. The number of characters and complexity of the story made it more difficult to talk about
every person while still being creative, but it allowed me not to fixate on the words I did not know which made it easier in that way.” (participant 24, complex task).

Another example that reflected the enhancing creativity was found in participant 36:

“I knew more of the vocabulary and was more familiar with what I was doing so I felt more comfortable and less judged when I stumbled. There was a lot more happening in the task with characters and storyline, etc. but it was good because I felt like I was able to say a lot more about what was happening, even if I forgot one word I could overcompensate by talking about other things in each picture.” (participant 36, complex task).

Supporting quantitative results (the self-assessed ratings), the current study found no statistical differences when comparing participants’ responses between the simple and the complex task. Specifically, results revealed that 50% of the participants—26 participants (out of 51) (half of the sample size)—clearly expressed that the complex task was experienced as more difficult when compared to the simple task. However, 21% of participants—41% of the total sample size—either did not explicitly compared the two tasks or found them both equally difficult. Those were categorized as “does not explicitly say”. Finally, 8% of participants believed that the simple task had been more difficult that the complex task. This group of participants reported that the task that posed higher cognitive demands seemed easier for them, as it allowed more room for creativity due to the greater amount of action happening in the pictures. Results from the qualitative data demonstrated that, although 50% of the participants overall (a significant amount) found the complex task to be more difficult, still the other half did not, and therefore, no statistical
differences were found for the task difficulty measure in the present dissertation.

Results revealed within the latter category ‘other factors’, the most relevant sub-categories mentioned by participants as factors influencing the degree of difficulty for the two tasks were, in order of importance: the presence or absence of planning time and/or the possibility of note-taking prior to task performance reported in 28% of the comments; time pressure in 26% of the total number of instances; task mode (speaking versus a written production task) and task type (grouped into the same sub-category) appeared in 22 % of the cases; task repetition and task familiarity were observed in a 5% of the data; and perceived level of proficiency in Spanish in 2% of the data.

As displayed in the coding frame above, within the ‘other factors’ category, most of the sources for task difficulty mentioned by participants were task-internal or task-external factors. For task-internal, these categories were mainly related to task complexity and task conditions, such as, the non-interactive nature of the tasks, the oral mode of the tasks, time pressure, and other contextual factors surrounding the tasks. In relation to task-external factors were factors such as, the absence of feedback from an instructor, the presence of a recording device that timed participants’ performance, the impossibility to write down ideas or thoughts prior to speaking, or not being provided with a source of vocabulary to guide them throughout in the task.

Many of these factors have been conceptualized in previous task complexity literature (e.g. Robinson, 2001; Skehan, 1996) and were related to task complexity within the resource-depleting dimension (Robinson, 2005)—such as the presence/absence of planning time and prior preparation—or certain task condition elements, such as task mode/modality, time pressure, task type, task mode (oral versus written mode), task
repetition.

As part of d) how and what aspects of performance were affected in relation to task difficulty, task performance, the difficulty of attending to multiple cognitive processes taking place simultaneously was another prominent category. In particular, for the complex task, 23% of participants’ comments claimed that the complex task (task B) was more difficult because it entailed performing too many cognitive activities to be performed all at once.

5.1.3.2 Mental Effort

For mental effort, data from 5 participants were lost, and thus, a total of 46 participants’ comments for each task were analyzed for this cognitive load measure.

Overall, results showed that there was a much clearer pattern between the two levels of task complexity regarding the amount of brain power exerted in each task. In general, comparisons indicated a higher percentage of agreement and less variability as compared to task difficulty.

For the qualitative data, results revealed statistically significant differences between the simple and the complex task regarding the amount of mental effort put into the task as perceived by learners. A total number of 28 participants (out of 46)—60% of the sample size—demonstrated that learners exerted a significantly higher amount of brain power in the complex task as compared to the simple task. These results supported the quantitative results presented in the previous section also indicating a meaningful increase in the mental effort employed by participants in the task that was designed to impose higher cognitive load.

Interestingly, in the majority of the cases, results showed that participants often
compared the two tasks regarding their task difficulty and the mental effort exerted:

Example: “This required more mental effort because I had to think harder about the words I used. I had to be more specific in my word choice and dig further down. I couldn’t just use the words on the top of my mind because the theme was more specific.” (participant 28, complex task).

As seen in the previous example, participant 28 described the amount of brain power put into the task by comparing it to the simple task. Such examples showed participants reflecting on their mental effort on a task with intended higher levels of cognitive load actual as compared to a task designed with lower levels of cognitive load. Let us consider a few representative comments extracted from the participants’ responses from the open-ended questions:

“Because there were more characters, background information, and action going on in each picture it took more mental effort to figure out what was going on in each picture as well as the entire story and then figuring out how to retell the story, so it made sense.” (participant 24, complex task).

“The task required a lot of mental effort because of the action taking place in the story. There were a lot of things going on that you had to process.” (participant 34, complex task)

“This task required me to use more mental effort than the last task, as I had to describe more.” (participant 51, complex task)

The excerpts just presented revealed how one of the factors that contributed the most to mental effort was again the main task complexity operationalization employed in the current dissertation, +/- Few Elements. Overall, the higher number of characters, and
the amount of action going on was reported by 32% of the cases for both the simple and the complex tasks.

In order of importance, however, the factor that was reported as major contributor to increase in the amount of mental effort put into the task was task-linguistic demands in 63% of cases. As for task difficulty, within the category of task-linguistics demands, data showed “the vocabulary elicited” in the cognitively complex task led to a meaningful increase in the mental effort.

As expected, data revealed that the mental effort exerted in both tasks was perceived by participants as a willing and conscious effort, thus, learner-dependent to engage in the task cognitively, rather than task-dependent. In this respect, participants’ comments provided evidence that mental effort, although closely connected, should be regarded as an independent construct different in nature from cognitive load. The following example reflected this difference:

“I started to get discouraged when I didn't know how to describe what I saw so I stopped putting in as much effort as the task went on.” (participant 22)

Data demonstrated this conscious and willing nature of mental effort, as opposed to the idea of cognitive load as inherent to the task, which refers to the actual cognitive demands of the task (task-dependent). This contrast between the two constructs is reflected in examples like this one:

“The reason why I did not put full mental effort in is because at a certain point I realized I just was not going to be able to explain the story since I knew so few relevant words, and I was not going to experience any consequences for stopping or acknowledging I failed. It required a lot of effort though because there were so many elements to consider,
Participant 8 acknowledged that the intended-to-be cognitively complex task indeed posed inherently higher cognitive demands, but he consciously decided not to engage mentally nonetheless.

Overall, 28% of participants considered that the speaking nature of the task (task modality/mode) was a contributor to the mental effort put into the tasks. 19% discussed that the absence of planning or preparation prior to task performance led to a greater amount of mental effort exerted in the complex task. 15% of participants reported that having done a similar task twice (known as task-repetition, Bygate 2001) influenced the amount of mental effort put into the tasks. Finally, time pressure or ‘feeling rushed’ was pointed out by 10% of participants in both the simple and the complex task.

5.1.4 Summary of Results for Research Question 1

Following recent TBLT researchers (e.g. Sasayama, 2016; Révész et. al, 2016), RQ1 examined the relationship between cognitive load—as measured by self-assessed task difficulty, self-assessed mental effort, and time estimation—and intended cognitive task complexity. Through the use of these subjective cognitive load measures, the current study sought to: 1) validate the cognitive demands of the task complexity to be able to test for the Anxiety - Cognition – L2 Task Performance relationship; and 2) contribute to the recent body of task complexity literature to examine the effectiveness of these cognitive load measures in Spanish as a L2.

The results above revealed that, for these particular groups of learners and low level of proficiency, out of the three subjective cognitive load measures, only mental effort seemed to have been an effective instrument in addressing the cognitive demands of the
two levels of task complexity—the designed-to-be simple and the designed-to-be complex tasks in the current study. For this measure, the cognitively complex task led to a significantly higher level of self-assessed mental effort exerted. In other words, participants overall perceived their brain power to be significantly higher in the complex narrative. This served to validate that the two main tasks employed in the current study indeed imposed significantly different cognitive demands on participants.

On the contrary, results demonstrated that task difficulty and time estimation did not seem efficient in distinguishing between the two levels of cognitive task complexity, at least for these particular pedagogic tasks. Interestingly, overall qualitative results supported quantitative results in the relationships observed. As demonstrated in the percentages for the various theme categories, a clear pattern shed light on better explicating the constructs of task difficulty and mental effort.

Results also yielded that the use of +/− few elements (as the main operationalization for cognitive task complexity) significantly contributed to increase participants’ levels of self-assessed task difficulty and mental effort. Importantly, vocabulary was revealed as the most important factor influencing task difficulty and mental effort.

5.2 Results for State Anxiety

5.2.1 Reliability Analyses for the Language Anxiety Questionnaires

RQ2a sought to find the relationship between cognitive task complexity (complex versus simple) and perceived levels of state anxiety of L2 Spanish learners during task performance. In order to answer the question, and before exploring descriptive statistics, reliability analyses were calculated with two purposes. First, to find out about the two anxiety instruments’ internal consistency (or the intra-class correlation coefficient) to
capture that all items of the scales measured the same constructs in the current study; and second, with the purpose of examining how consistent participants’ responses were for the state anxiety measure over time (at Time 1 and 2), and also across the two task complexity levels (the cognitively simple versus the cognitively complex task).

Although the FLCAS was maintained as the original version from Horwitz (1986), the state anxiety scale was adapted to the current study, and new items were included, adjusted and excluded. Therefore, reliability analyses were specifically important for the state anxiety questionnaire to ensure the instrument’s effectiveness in capturing the levels of state anxiety for the two cognitive complexity levels tasks. As highlighted earlier, exploring state anxiety at two different points in time throughout the task would shed light on the dynamicity of emotional states, such as language anxiety, and how they fluctuated over time. For the state anxiety scale, Cronbach’s alpha revealed to be acceptable when comparing participants’ scores between the simple and the complex task after the first administration in Time 1 (0.82) and (0.74) after the second administration in Time 2. This indicated that all items of the questionnaire were effective at tapping the anxiety construct, being responses not randomly provided by participants.

Cronbach estimates (0.93) when comparing participants’ scores between Time 1 and Time 2 for the simple task at Time 1 were excellent, and (0.92) between Time 1 and Time 2 for the complex level of cognitive task complexity. These results demonstrated items of the construct and participants’ responses were highly consistent in distinguishing among participants over time, from the mid-task to the post-task administration. Lastly, the Cronbach’s alpha estimates for the classroom anxiety scale, FLCAS, indicated high internal consistency in distinguishing among participants (0.81), again showing the
instrument was reliable at measuring participants’ classroom anxiety in Spanish as a second language.

5.2.2 Quantitative Results for Perceived State Anxiety: Simple and Complex Tasks in Time 1 and Time 2

RQ2a sought to examine how cognitive task complexity was connected to levels of state anxiety (as perceived by L2 learners) during oral task performance in L2 Spanish. To respond to the question, descriptive statistics, were calculated for participants’ state anxiety scores on the cognitively simple and the complex task.

Data was inspected for outliers, and this time, no severe univariate or bivariate outliers were identified. Participants 14, 39 (univariate outliers) and 29 (bivariate outlier) scored only slightly below the lower bound 95% CI. However, these were not considered extreme outliers, and therefore, they were kept in the sample data set. Distributions of normality and histograms were explored for levels of state anxiety for the two task complexity levels (RQ2a), and from Time 1 to Time 2 (RQ3). For both the simple and the complex task, and half-way through the task (Time 1) and right after task completion (Time 2), again no severe skewness or kurtosis was revealed, and descriptive statistics showed reasonably normally distributed data sets in all cases.

5.2.3 Results on the Fluctuation of State Anxiety Over Time (From Time 1 to Time 2) and its Dynamicity on the Simple Task

The section below presents on the results addressing RQ3 on the fluctuating nature of state anxiety as measured at two-points in time throughout the task: half-way through the task and after task completion. For the cognitively simple task, state anxiety scores at Time 1, as captured during task performance, ranged between 1.70 (minimum score) to
5.50 (maximum score) out of a possible total score of 6 in the Likert scale. Additionally, state anxiety scores spread over a range of 3.80, with the mean situated in the 3.89 (see Table 14). Although minimal negative skewness was observed, with slightly more participants being placed towards the higher range of the scale, scores were normally distributed (see Figure 15).

**Table 14**: Descriptive statistics for state anxiety for the simple task at Time 1.

<table>
<thead>
<tr>
<th></th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>51</td>
</tr>
<tr>
<td>Valid</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>3.89</td>
</tr>
<tr>
<td>Trimmed Mean</td>
<td>3.92</td>
</tr>
<tr>
<td>95% Confidence Interval</td>
<td></td>
</tr>
<tr>
<td>for Mean</td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>3.64</td>
</tr>
<tr>
<td>Upper bound</td>
<td>4.14</td>
</tr>
<tr>
<td>Median</td>
<td>3.90</td>
</tr>
<tr>
<td>Variance</td>
<td>0.78</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.88</td>
</tr>
<tr>
<td>Minimum</td>
<td>1.70</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.50</td>
</tr>
<tr>
<td>Range</td>
<td>3.80</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.58</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.02</td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td>0.33</td>
</tr>
<tr>
<td>Std. Error of Kurtosis</td>
<td>0.65</td>
</tr>
</tbody>
</table>
On the simple task, for the levels of state anxiety after task completion (Time 2), the mean scores showed statistically significant differences between Time 1 and Time 2. This provided evidence that participants perceived their state anxiety levels to be significantly higher after completing the task ($M= 4.14$) than while they were performing it ($M= 3.89$). In this case, scores spread over the same range as at Time 1 (3.80), from a minimum score of 2.10 and a maximum of 5.90 (see Table 15). As for Time 1, it was concluded that distributions of scores were normal, only showing minimal negative skewness (see Figure 16).
Table 15: Descriptive statistics for state anxiety for the simple task at Time 2.

<table>
<thead>
<tr>
<th></th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Valid 36</td>
</tr>
<tr>
<td></td>
<td>Missing 0</td>
</tr>
<tr>
<td>Mean</td>
<td>4.14</td>
</tr>
<tr>
<td>Trimmed Mean</td>
<td>4.15</td>
</tr>
<tr>
<td>95% Confidence Interval</td>
<td>Lower bound 3.90</td>
</tr>
<tr>
<td>for Mean</td>
<td>Upper bound 4.37</td>
</tr>
<tr>
<td>Median</td>
<td>4.20</td>
</tr>
<tr>
<td>Variance</td>
<td>0.68</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.82</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.10</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.90</td>
</tr>
<tr>
<td>Range</td>
<td>3.80</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.44</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.20</td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td>0.33</td>
</tr>
<tr>
<td>Std. Error of Kurtosis</td>
<td>0.65</td>
</tr>
</tbody>
</table>
For the simple task, since more participants were placed in the right-hand side of the scale, overall the majority of participants perceived their levels of anxiety to be higher than lower (the majority of scores were above 3). In particular, a significantly higher number of participants “slightly agreed” and “agreed” that the task “was being” (Time 1) or “had been stressful” (judging retrospectively for Time 2). For this particular population of learners, even a task “design-to-be-simple” was reported to induce quite elevated levels of language anxiety (on a scale 1-6), and this anxiety significantly increased over time. In other words, perceived levels of state anxiety fluctuated from Time 1 to Time 2, being these differences statistically significant. As predicted, the changing nature of language anxiety was evidenced on the different levels of state anxiety captured at different points during task performance. Additionally, as it occurred in the pilot study, state anxiety suffered a
meaningful increase from Time 1 ($M = 3.89$) to Time 2 ($M = 4.14$) (rather than a decrease).

5.2.4 Results on the Fluctuation of State Anxiety Over Time (From Time 1 to Time 2) and its Dynamicity on the Complex Task

A very similar pattern was discovered when looking at the scenario of the cognitively complex task (RQ3). At Time 1 (half-way through task performance), descriptive statistics and graphical representations were conducted, revealing reasonably normal distributions of scores. As illustrated in Figure 17, no severe skewness or kurtosis was observed. As for the simple task, descriptive statistics (Table 16) depicted the same results for the complex task. The negative skewed value (-0.13) did not exceed twice the standard error of skewness (0.66), confirming the skewness of the data was only minimal.

![Figure 17: Histogram for levels of state anxiety for the complex task at Time 1.](image)

Following the pattern of the simple task, for the complex task, levels of state
anxiety—as perceived by participants—fluctuated significantly over time. Descriptive statistics provided evidence that statistically significant differences existed between Time 1 ($M = 4.29$) and Time 2 ($M = 4.40$) for state anxiety during task performance for the cognitively complex task. Between half-way through the task (while the task was still in progress) until its completion (post-task administration), participants believed their anxiety levels to increase significantly over Time.

5.2.5 On the Relationship between Cognitive Task Complexity and State Anxiety at Time 1 and Time 2

For the complex task, a meaningful increase occurred in the perceived levels of state anxiety when compared to the simple cognitive complexity condition. The minimum and maximum scores were of 2.50 and 5.70, respectively; and the mean scores for the complex task at Time 1 was statistically higher ($M = 4.29$) as compared to the mean score for simple task at Time 1 ($M = 3.89$) (see Table 14 & 16). Confidence intervals in Figure 19 display the same information visually. These results confirmed the current study’s hypothesis that increasing the cognitive load of the task would lead to higher levels of state anxiety (as perceived by participants), and these differences were meaningful.
Table 16: Descriptive statistics for state anxiety for the complex task at Time 1.

<table>
<thead>
<tr>
<th></th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Trimmed Mean</td>
<td></td>
</tr>
<tr>
<td>95% Confidence Interval</td>
<td>Lower bound</td>
</tr>
<tr>
<td>for Mean</td>
<td>Upper bound</td>
</tr>
<tr>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
</tr>
<tr>
<td>Std. Deviation</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td></td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td></td>
</tr>
<tr>
<td>Std. Error of Kurtosis</td>
<td></td>
</tr>
</tbody>
</table>

When examining the complex task at Time 2, Figure 18, again, revealed only minimal negative skewness in the distributions of scores. As for Time 1, a similar pattern was revealed, where a meaningful increase in the mean scores was observed between the cognitively simple ($M = 4.14$) and the cognitively complex task ($M = 4.40$) at Time 2. Hence, a statistically significant effect was demonstrated for perceived state anxiety between the task that posed lower cognitive demands and the task that posed higher cognitive task demands at Time 2 (see Tables 15 & 17, and Figure 20).
Table 17: Descriptive statistics for state anxiety for the complex task at Time 2.

<table>
<thead>
<tr>
<th>Total score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$ Valid</td>
<td>51</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>4.40</td>
</tr>
<tr>
<td>Trimmed Mean</td>
<td>4.41</td>
</tr>
<tr>
<td>95% Confidence Interval for Mean</td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>4.21</td>
</tr>
<tr>
<td>Upper bound</td>
<td>4.58</td>
</tr>
<tr>
<td>Median</td>
<td>4.40</td>
</tr>
<tr>
<td>Variance</td>
<td>0.43</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.65</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.90</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.90</td>
</tr>
<tr>
<td>Range</td>
<td>3.00</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.29</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td>0.33</td>
</tr>
<tr>
<td>Std. Error of Kurtosis</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Importantly, the same statistical effect in participants’ state anxiety was revealed between the simple task at Time 1 and the complex time at Time 2. However, only a marginal effect seemed to occur between the simple task at Time 2 and the complex task at Time 1.

The above results also suggested an interaction between cognitive task complexity and Time. This means that differences across task conditions were significant over time (from Time 1 to Time 2). This potential interaction was further examined by a One-way Repeated Measures ANOVA conducted using two-within factors. Results for the inferential analyses will be discussed in the following section.

5.2.5.1 Inferential Analyses for State Anxiety at Time 1 and Time 2

To confirm the results observed on the descriptive statistics, One-way Repeated
Measures ANOVA tests were conducted separately on each level of cognitive task complexity and at both times throughout task performance. Overall within-group differences on the perceived state anxiety levels for the cognitively simple and complex task were investigated, using task complexity level as the repeated condition. These analyses allowed the researcher to explore the effect of task complexity alone on state anxiety. After that, within-group comparisons were performed separately for each task at Time 1 and Time 2 to investigate the effect of Time on state anxiety.

As expected, at Time 1, statistical significant differences were confirmed between the cognitively simple and the cognitively complex task for perceived state anxiety ($N = 51$) $F (1,50) = 17.90$, $p = 0.000$, partial Eta squared = 0.264. At Time 2, within-group comparisons showed the same pattern, revealing a statistical significant effect for cognitive task complexity on levels of state anxiety ($N = 51$) $F (1,50) = 7.51$, $p = 0.008$, partial Eta squared = 0.131. Cohen’s (1988) revealed a large effect size ($d = 0.62$) for cognitive task complexity at Time 1; and a medium effect size ($d = 0.39$) at Time 2.

The statistical differences for task complexity on levels of state anxiety are also displayed in Figures 19 and 20 for Time 1 and 2, respectively, where the mean score for the complex task exceeded the upper confidence interval of the simple task.
Figure 19: Error bars with confidence intervals for state anxiety at Time 1 based on the two cognitive complexity levels.
One-Way Repeated Measures ANOVAs were conducted once again separately on each task employing Time as the within-group factor to see if time made a difference in how levels of task anxiety were perceived by participants.

Confirming previous results on the descriptive statistics, inferential analyses showed that, for the simple task, state anxiety fluctuated in a meaningful way as time passed during task performance, from Time 1 (half-way through the task) to Time 2 (right after the task) ($N = 51$) $F(1,50) = 17.09$, $p = 0.000$, partial Eta squared = 0.255. For the cognitively complex task, within-group comparisons also demonstrated statistical

---

**Figure 20**: Error bars with confidence intervals for state anxiety at Time 2 based on the two cognitive complexity levels.
significant differences between Time 1 and Time 2 on learners’ state anxiety \((N = 51)\) \(F(1,50) = 4.25, p = 0.04\), partial Eta squared = 0.078. Cohen’s (1988) revealed a medium effect size for Time for the simple task \((d = 0.57)\); and a small to medium effect size \((d = 0.28)\) for the complex task.

Figures 21 and 22 provide the same information regarding the effect size visually for Time when examining the confidence intervals between Time 1 and Time 2 on the two task complexity levels.

**Figure 21:** Errors bars for the simple task exploring state anxiety over time (Time 1 versus Time 2).
Lastly, a potential interaction between cognitive task complexity and Time on levels of state anxiety was examined. In order to see if an interaction existed between the two independent variables, a two-within factor Repeated Measures ANOVA was performed with task complexity as the first within-subject factor, and Time as the second within-subject factor.

This analysis was deemed important as it considered the effects of the two independent variables together on the main dependent variable of interest in this research question (state anxiety). Once again, a main effect was revealed for task complexity \((N = 51)\) \(F (1,50) = 14.30, p = 0.000\), partial Eta squared = 0.22; and for Time \((N = 51)\) \(F (1,50) = \ldots\)
= 18.67, \( p = 0.000 \), partial Eta squared = 0.27. Partial eta squared of 0.22\(^7\) and 0.27 indicated a large effect size. Interestingly, results of this analysis showed an interaction effect also took place between cognitive task complexity and Time on state anxiety (\( N = 51 \)) \( F (1,50) = 3.73, p = 0.05 \), partial Eta squared = 0.070, with a medium size effect. This interaction effect was also observed in the graph in Figure 23 below:

![Profile plot for the interaction between cognitive task complexity (simple versus complex) and Time (Time 1 versus Time 2) on state anxiety.](image)

**Figure 23:** Profile plot for the interaction between cognitive task complexity (simple versus complex) and Time (Time 1 versus Time 2) on state anxiety.

\(^7\) A partial eta squared between 0.14-0.20 is considered a large effect size, between 0.06 and 0.11 a medium effect size; and between 0.010 and 0.039 a small effect size.
5.2.6. Testing the Effect of Order of Tasks on State Anxiety at Time 1 and Time 2

To ensure the order of the tasks did not play a role in participants’ self-assessed scores on state anxiety for the two levels of task complexity, the effect of order of text was tested for both Time 1 and Time 2. All participants that received the complex task first and the simple task in the second place (complex first – simple second) were grouped together and compared to those participants that performed the simple task first and the complex task afterwards (simple first – complex second).

For Time 1, descriptive statistics (see Table 18) showed that performing the complex task first (complex first – simple second) had an advantage for participants on the cognitively simple complexity level, since state anxiety levels were lower ($M = 3.64$) than for the reversed order ($M = 4.13$). On the contrary, when participants performed the cognitively complex task in the second place, right after the simple task (simple first - complex second), the same benefit was not observed. In this case, the average for the simple task was statistically higher, indicating that the complex task seemed to have prepared participants slightly more, helping ease their language anxiety during task performance. For the complex task on Time 1, the scenario was similar, means and SDs showing a slight increase for the cognitively simple task. However, differences in their perceived anxiety were not statistically significant in this case. These results appeared to reveal that the order of tasks did not make a difference for the complex task.
Table 18: Descriptive statistics for the order of tasks on state anxiety for the two levels of task complexity at Time 1.

<table>
<thead>
<tr>
<th>Order of tasks</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S - C</td>
<td>C - S</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>4.13</td>
<td>3.64</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.85</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>Min</strong></td>
<td>2.10</td>
<td>1.70</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>5.40</td>
<td>5.50</td>
</tr>
<tr>
<td><strong>95% Confidence Interval</strong></td>
<td>3.78</td>
<td>3.28</td>
</tr>
<tr>
<td><strong>Lower bound</strong></td>
<td>4.47</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Upper bound</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To confirm whether the order of the two tasks played a role on participants’ perceptions of their levels of state anxiety, one-way repeated measures ANOVA were performed separately for Time 1 and Time 2. As all participants performed took part in the two story-telling, for the analyses the order of tasks was treated as between – subject’s factor and the cognitive task complexity level as the within – group factor. A statistically significant interaction between order of tasks and cognitive task complexity (order of tasks*task complexity) corroborated the descriptive statistics indicating a statistically significant effect for order of tasks on state anxiety at Time 1 \((N = 51)\) \(F(1,49) = 4.67, p = 0.03\), partial Eta squared = 0.08, and at Time 2 \((N = 51)\) \(F(1,49) = 5.41, p = 0.02\), partial Eta squared = 0.10.
The steeper blue line for the simple task, displayed on Figure 24 showed a statistically significant decrease on self-assessed language state anxiety between the two task orders. On the contrary, state anxiety did not seem to be equally affected by the order of the task on the complex task (green line).

![Figure 24](image.png)

**Figure 24:** Profile plots for the interaction between the order of tasks and cognitive task complexity (simple versus complex) on language state anxiety at Time 1.

For Time 2, a similar pattern was revealed when the order effect of the two stories was tested after participants’ task completion. Participants’ perceived their language anxiety levels to be lower on the simple task ($M = 3.90$) when the complex task was
performed in the first place (C-M) as compared to when it was administered in the second place ($M = 4.36$) (see Table 19 & Figure 25).

**Table 19:** Descriptive statistics for the order of tasks on state anxiety for the two levels of task complexity at Time 2.

<table>
<thead>
<tr>
<th></th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State Anxiety</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order of tasks</td>
<td>S - C</td>
<td>C – S</td>
</tr>
<tr>
<td></td>
<td>S - C</td>
<td>C – S</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>4.36</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>4.41</td>
<td>4.37</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>Min</strong></td>
<td>2.70</td>
<td>2.10</td>
</tr>
<tr>
<td></td>
<td>2.90</td>
<td>3.00</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>5.90</td>
<td>5.30</td>
</tr>
<tr>
<td></td>
<td>5.90</td>
<td>5.50</td>
</tr>
<tr>
<td><strong>95% Confidence</strong></td>
<td>4.06</td>
<td>3.54</td>
</tr>
<tr>
<td><strong>Interval</strong></td>
<td>4.11</td>
<td>4.14</td>
</tr>
<tr>
<td><strong>Lower bound</strong></td>
<td>4.66</td>
<td>4.25</td>
</tr>
<tr>
<td><strong>Upper bound</strong></td>
<td>4.77</td>
<td>4.70</td>
</tr>
</tbody>
</table>
Figure 25: Profile plots for the interaction between the order of tasks and cognitive task complexity (simple versus complex) on language state anxiety at Time 2.
5.2.7 Qualitative Data: Open-Ended Questions for State Anxiety on the Simple and Complex Task

To shed light on the relationship between cognitive task complexity and state anxiety (RQ2a) as perceived by participants, percentages on each category were calculated for the open-ended questions. Participants’ responses revealed important findings regarding the relationship between cognitive task complexity and perceived state anxiety. Although some participants provided a short comment of one or two sentences, some of them wrote a longer and more detailed explanation of why and how their state anxiety and performance were affected in the two task complexity levels. In what follows, results in relation to the qualitative data are discussed.

5.2.7.1 Emotions


Many of these adjectives and nouns were also mentioned in connection with task difficulty; for example: “Simplicity made the task much easier than the last comic. However, the objects which I didn't know the vocab words for stressed me out and made the task more difficult than it had been for the first few frames.” (simple task, participant 9).

This was interesting as it confirms how language anxiety is strongly related with the task difficulty component in Robinson’s triadic model (2001), as one of the individual differences that determines the perception of difficulty of L2 tasks.

The fluctuating nature of language anxiety and its unpredictable behavior in relation to different cognitive task demands was revealed in the variety of participants’ comments. In terms of the differences between the cognitively simple and the cognitively complex task, 29 participants out of 51 (57%) believed their anxiety levels were lower in the simplest task as compared to the cognitively complex task. These same participants thought they also performed better in the simpler task overall. On the contrary, 8 participants out of 51 (23%) considered that their performance in the more complex task was better and they also felt less anxious and more comfortable doing it. Finally, 10 participants (19%) did not explicitly compare their performance in their tasks, did not consider one task was easier or more difficult than their other, or the easier aspects balanced out the more difficult ones. Those participants judged the two tasks to be equally difficult and anxiety-provoking.

5.2.7.2 Facilitative Versus Debilitative Language Anxiety

Interestingly, for the majority of those 57% of the participants who experienced the
complex task as more stressful, their performance was overall negatively affected, and the anxiety considered mostly a debilitative type of anxiety. The majority of the participants expressed that the cognitively complex task—mainly due to the high number of elements and action present in the story—was felt as extremely overwhelming inducing state anxiety and thus, leading to poor performance. Therefore, results indicated that for these participants, in particular, state anxiety had a debilitative role in task performance in both the simple and the complex task. As predicted, the detrimental effect was especially more acute when learners performed the cognitively complex task over the simple task. The following extract shows how increasing the number of elements of the story (+/- characters) function as a trigger for state anxiety (or debilitative anxiety) leading to poor task performance.

“When there are more elements to a story it just becomes more stressful and harder to talk about in Spanish. For example, la policía is a simple word in Spanish but I could not remember how to say police, just because I was stressed about all the aspects of the story.” (participant 1, complex task)

On the other side of the spectrum, for 11 participants (21%), state anxiety had a facilitative effect in both tasks’ performance, turning out to be a positive, fun, engaging and enjoyable experience. When anxiety was felt as facilitative, participants perceived their performance in the complex narrative task to be significantly better than their performance in the simple task. The majority of these participants attributed the facilitative role of anxiety to the advantage of having a wider array of characters and features in the story they could focus on. To this end, participants highlighted that having a greater variety of alternatives with regards the number of features they could pay attention to in a particular
photo allowed them to be more creative in the language elicited. Often, whenever participants did not have the linguistic resources to describe a particular action in the story, they could focus on another feature present in the photo they could describe in an easier way, which allowed them to communicate in the L2 and still achieve the task goal.

“It was also less anxiety-provoking when there was more happening in the picture, such as the first image, because then I could describe whatever I wanted to, rather than in the second image, where there is one specific thing happening that I need to know how to describe.” (participant 8, complex task)

Overall, 41% of participants used positive terms as psychological coping strategies to refer to the simple task, and 25% for the complex task. Among the most widely used, high frequency positive emotions were: feeling ‘good’, ‘confident’, ‘comfortable’, ‘satisfied’, ‘fun’, ‘encouraged’, ‘calmed’ and ‘enjoyment’. Other less frequently employed emotional terms were ‘joy’, ‘happy’, ‘prepared’, ‘excited’, ‘motivated’, and ‘humorous’. Overall, when the complex task was described positively, it was ‘motivating’, ‘fun’, ‘humorous’, ‘open for creativity’, and with ‘more room to narrate’, whereas the simple task was described positively in terms of being less overwhelming and stressful for participants because it contained less action and fewer elements.

5.2.7.3 Task-Internal Features

For the complex task, following Sasayama (2016), results showed that for 30 (out of 51) or 58% of the participants, the higher ‘number of elements’ (characters) contained in storyline contributed to debilitating state anxiety at times during task performance. Participants often referred to this as the higher number of characters or as the amount of action contributing to their state anxiety by feeling cognitively overwhelmed and anxious
to be able to describe so much information in so little time. For these participants, this task feature clearly made a difference in relation to the cognitive demands imposed between the two task complexity levels. Although initially, storyline and number of characters were coded separately, for the analyses they were conflated into the same feature. The rationale behind this decision was that participants often seemed to refer to the same task feature when described the amount of action in the storyline and the number of characters or elements contained.

On the contrary, for approximately 13 participants (25%) of the participants, the more complex storyline was felt as more “engaging” and “interesting” since there were more elements and more action involved. Overall, for these participants, having a more cognitively complex storyline seemed beneficial since they perceived they had “more to talk about” and felt less anxious about the task. Lastly, 8 participants (15%) of the participants did not mention the number of elements or the storyline as being either an advantage or disadvantage.

Clarity of pictures was mentioned by only 3 participants (5%) of participants for the complex task as extraneous cognitive load (Sweller, 1994). Three participants out of fifty-one indicated that in some of the pictures, distinguishing among the several characters and the obscure or the fuzzy elements caused confusion and made them take longer to figure out what was happening in the story. Finally, the length of the task was mentioned by only 2 participants (3%) as a problematic issue in relation to the complex task, described as taking longer due to the amount of action contained. However, due to its low-frequent appearance in the data, it did not appear to be a problematic element having any major impact on their perceived state anxiety, task difficulty or linguistic performance.
5.2.7.4 Task-Linguistic Demands

5.2.7.4.1 Vocabulary

Overall, vocabulary was perceived by participants as the most important source of debilitative state anxiety, believed to have affected task performance either for the simple or the complex task for all participants (100%) of participants. Following Skehan’s conceptualization of code complexity (1996, 1998), vocabulary was the most mentioned task-internal factor for both the simple and the complex task regarding code complexity. Specifically, for the simple task, ‘not remembering’ or ‘not knowing’ the right vocabulary was identified also as the most anxiety-provoking element in 84% (43 participants out of 51) of the cases. The remaining 16% of participants believed vocabulary was not difficult or it was not mentioned. Particularly, “unknown” words such as “alarm”, “pillow”, “light”, “fall”, “lamp” and “blinds” were constantly reported as a source of state anxiety, affecting task performance. In many cases, participants highlighted their frustration when realizing they could not narrate such a basic story since they were missing the key words. For the complex task, vocabulary was reported to be detrimental to task performance at an even higher degree, 96% of participants (49 out of 51). The most problematic words were: “ladder”, “leash”, “sidewalk”, “jail”, “hit”, “arrest”, and “handcuffs”. For the remaining 4% of participants, vocabulary in the complex task was not a viewed as a problem or it was not mentioned whatsoever.

5.2.7.4.2 Grammar

To a lesser degree, but still a high frequency stressor, was grammar being reported by 65% of participants (33 participants out of 51) overall for the two tasks. For the cognitively simple task, 31% (16 participants out of 51) of participants as a major
contributor to state anxiety; and for the cognitively complex task, 54% of participants (28 participants) reported grammar, the use of conjugations, and verb tenses accurately as a problematic feature. A majority of participants (69%) did not report ‘grammar’ as a trigger for language anxiety during performance on the simple task. However, for the complex task, a different pattern was revealed where 23% of participants did not perceive the grammar elicited in the task was detrimental in relation to state anxiety.

Not having knowledge of expressions such as “to throw a pillow”, “to fall”, and “to look out the window” was perceived as one of the most troublesome issues regarding code complexity for the simple task. For the complex task, grammatical expressions such as “to walk the dog”, “running away from” “to go to prison”, “to go to the hospital”, “to drive”, and “swerving” turned out to be the most problematic structures.

For the two tasks, participants mentioned verb conjugations and tenses to be to be the most difficult and anxiety-provoking factors. In particular, confusion about going back and forth between the present and past tense in narrating the story was highlighted by many participants.

Finally, with regard to ‘other factors’, the following elements were perceived by participants to affect their language state anxiety and L2 performance. First, a) ‘speaking anxiety/task mode’ was mentioned by 27% of participants for the simple task and 31% for the complex task; b) the difficulty in not having the opportunity to plan/prepare, ‘lack of planning/preparation’ or no-note taking for the task was mentioned by 11% of participants for the simple task and 17% for the complex task; c) overall, 21% of participants commented on the beneficial effects that performing the task a second time, namely ‘task repetition’ had on their language anxiety; d) ‘feeling rushed’ or a ‘lack of sufficient time’
was reported by 11% of participants for the simple and 9% for the complex task; e) task-type or the preference for an interactive or another type of task was revealed in the data only by 13% in the two tasks as a factor influencing state anxiety and task performance; f) task familiarity as the task progressed was mentioned by 5% in the complex task; g) fear of negative evaluation (Horwitz et al., 1986; Young, 1991) triggered by the presence of a native-speaker (primary investigator) and other Spanish L2 learners (participants participating in the study) (translated in a classroom context, the same feelings of apprehension are usually triggered by the presence of the professor and other classmates) was reported by 5% of participants; h) 3% of the participants mentioned feeling anxiety due to the distracting environment conditions for both the simple and the complex task, such as, being in the lab (versus the classroom), being recorded or the fact of speaking to a computer (versus speaking to a person).

In some cases, language state anxiety was reflected in some participants and revealed through “catastrophizing” (Grohol, 2018) (a common cognitive indicator of state anxiety) tendencies. Let us consider the following excerpt:

“During this task I felt very stressed and slightly embarrassed because I could not recall simple vocabulary. I also began to feel worried that my poor performance in this task may mean that I will begin to have more difficulties in my Spanish class and that I will never be able to become fluent in Spanish.” (participant 17, simple task)

Catastrophizing was defined by Grohol (2018) as an irrational thought that human beings have, when they believe that a particular situation in the present is far worse than it actually is. This generally takes the form of imagining a catastrophe out of a current or a future situation by exaggeration of reality. In the previous example, the participant’s state
anxiety made him catastrophize by imagining a future failure that had not happened and was very unlikely to occur.

**5.2.8 Relationship Between Perceived State Anxiety and Classroom Anxiety in the Spanish Classroom**

As part of RQ2b, the current dissertation sought to investigate if a relationship existed between learners’ tendencies to experience language anxiety in the Spanish classroom and their perceived state anxiety during performance in the simple and the complex task. In order to answer explore the potential relationship, participants’ state anxiety scores at Time 1 and Time 2 were correlated to the FLCAS, used as a baseline for classroom anxiety in the Spanish classroom to better understand the role of participants’ state anxiety on L2 tasks.

Although a positive correlation was revealed between the two cognitive complexity levels and the classroom anxiety scores, this correlation was not statistically significant. In particular, Pearson’s correlations showed a very weak correlation between state anxiety and classroom anxiety scores (FLCAS) for the simple task at Time 1 \( r (51) = 0.10, p = 0.44 \), and also at Time 2 \( r (51) = 0.09, p = 0.51 \). A similar pattern was shown for the complex task where no relationship occurred between the FLCAS scores and state anxiety scores at Time 1 \( r (51) = 0.008, p = 0.95 \), and at Time 2 \( r (51) = -0.024, p = 0.86 \).

In the current study, classroom anxiety was used as baseline to achieve a better understanding between language anxiety—as an emotional state—and classroom anxiety—as a general predisposition to be anxious in the foreign language classroom. However, no relationship was found between classroom anxiety in L2 Spanish and state anxiety during task performance either for the cognitively simple or the cognitively
complex task. Results demonstrated a lack of relationship\textsuperscript{8} between these types of anxiety dimensions, revealing the independent nature of the two constructs in an instructed classroom setting.

5.2.9 Summary of Results for Research Question 2 and Research Question 3

RQ2a investigated the relationship between cognitive task complexity and levels of perceived state anxiety at two points in time during task performance, half-way through the task (Time 1) and after task completion (Time 2). Both quantitative and qualitative results revealed statistical significant differences between the simple and the complex task. A meaningful increase was revealed in relation to the levels of state anxiety for the cognitively complex task (as compared to the simple task) for Time 1 and Time 2. A medium effect size was observed in the self-assessment state anxiety measure between the cognitively simple task and the cognitively complex task at Time 1 and at Time 2.

In addition, qualitative results revealed interesting results on what factors contribute the most to induce learners’ state anxiety. These factors are mainly task-internal features and potentially other learner factors (individual differences such as personality, learning styles, etc.) that were not addressed in this study). For the purpose of the current study, the following trends have been revealed in relation to the open-ended questions:

First, when comparing the two tasks, 57% reported a much higher amount of negative words related for the designed-to-be complex task over the simple version.

\textsuperscript{8} The relationship between classroom anxiety and state anxiety during task performance was examined by dividing the sample into highly anxious and low anxious learners according to their general predisposition for foreign language anxiety. An attempt was made, based on participants’ FLCAS scores, to compare the two groups (High vs Low anxiety) for their perceived levels of state anxiety in both cognitive complexity levels. However, data proved inadequate for group division and, thus, analyses were discarded for further investigation.
Additionally, these participants usually made comparisons between the two tasks stating that the complex task had been much more stressful, overwhelming and difficult (as compared to the other task) and provided reasons that contributed to trigger anxiety during task performance. For these participants, anxiety was revealed as a negative and debilitative type of anxiety.

Second, among the task-linguistic demands, a lack of vocabulary was mentioned by 84% of the participants for the simple task, and by 96% for the complex task as the most anxiety-provoking factor; followed by grammar (31% for the simple and 54% for the complex task.

Third, for these participants, +/- few elements (+/- few characters) and the amount action in the complex task was considered as the most anxiety-provoking contributor with regards task design features by 58% of participants. These findings supported Sasayama (2016) in the operationalization of +/- Few elements as a factor making a difference to increase the actual cognitive demands in L2 tasks. In addition, these results also indicated that an increase in cognitive task complexity led to significantly higher levels of state anxiety.

In relation to ‘other factors’ that contribute to state anxiety, results revealed the following: task mode with 27% in the simple task and 31% for the complex task; and a lack of planning or preparation prior to the task, with 11% for the simple and 17% for the complex task. Lastly, task repetition was evidenced as another influential factor in the degree of state anxiety experienced by 21% of participants overall.

Finally, to a lower extent, other factors that were related to language anxiety in the current study had to do with task conditions or task-internal factors. Some of these triggers
were: time pressure, task familiarity, the presence of native speaker, and other environmental distractions (among others).

In the opposite line of results, 13 participants—25% of the total sample size—believed that they experienced less anxiety and felt more comfortable performing the complex task over the simple one. For these participants language anxiety was revealed as a facilitative type of anxiety, and the task seen as more engaging and interesting, promoting more quality and quantity of language. Lastly, 8 participants (15%) seemed to have experienced similar levels of language anxiety during the two tasks. These learners did not make explicit comparisons in their levels of anxiety or their overall performance in their tasks.

RQ2b sought to investigate if any relationship existed between participants’ classroom anxiety scores (as measured by Horwitz et al.’s, 1986 foreign language classroom anxiety scale, FLCAS), used as baseline to interpret state anxiety during both the simple and the complex time. Results indicated no relationship between the perceived state anxiety scores and participants’ general tendencies to experience language anxiety in the Spanish classroom.

Finally, quantitative and qualitative results supported that language anxiety (as perceived by L2 learners) is a dynamic construct that constantly fluctuates (RQ3), and it increases or decreases according to factors of different natures that are in continuous interaction.

Importantly, quantitative results also revealed an interaction between cognitive task complexity and Time on state anxiety. In other words, participants’ state anxiety levels varied across task complexity levels, and that variability was directly related to the timing.
when that anxiety was measured. To this end, state anxiety fluctuated from Time 1 to Time 2 significantly, and task complexity was a mediating factor in that relationship.

5.3 Results on the Relationship between Language Anxiety, Cognition and Task Performance in Second Language Spanish When Cognitive Task Complexity is Manipulated: The Three-Way Interface of Language Anxiety — Cognition — Second Language Task Performance

RQ4 was divided into as RQ4a and RQ4b. This question sought to shed light on how state anxiety (as perceived by L2 learners of Spanish) is related to cognitive processes (RQ4a) and to task performance (RQ4b) in L2 Spanish when cognitive task complexity is manipulated. The question was addressed by examining participants’ responses on the open-ended questions regarding how they believed their perceived state anxiety impacted their performance during both the simple and the complex task.

5.3.1 On the Relationship between Language State Anxiety and Cognitive Processes During Oral Task Performance when Cognitive Load is Manipulated (Research Question 4a)

To respond to RQ4a, I will report on the main cognitive functions that participants perceived were mainly affected by state anxiety in the two different levels of cognitive complexity.

5.3.1.1 Short-Term Memory/ Working Memory

As predicted, results confirmed the findings of the pilot study, indicating that learners perceived one of the most prominent brain functions to be directly affected in a negative way by language state anxiety was the short-term memory or working memory.

“I felt anxious, flustered, and nervous when completing this task because I am not
a strong Spanish speaker so anytime I have to speak in Spanish I get nervous and tend to forget the Spanish that I do know and I get stuck when I do not know the correct vocab words to describe things. I also felt this way because there were other people in the room while I was speaking.” (participant 6, complex task).

As seen in participant’s 6 subjective interpretation of the situation, feeling anxious and nervous seemed to impede temporarily the access to the learner’s long-term memory. Importantly, the participant recognizes that this phenomenon regularly happens to him whenever he is feeling anxiety.

5.3.1.2 General Thought Process

Another commonly reported cognitive function that seemed to have suffered from this emotional state, as reported by participants, was their normal “thought” process. In this respect, many participants referred to not being able to think normally and access the necessary linguistic resources already stored in their memory to express what they need.

“When I did not know how to say a word or how to phrase something, I panicked and could only think about how I could not say what I wanted to. I was unable to think of another way of saying something because I could only think of what I could not say.” (participant 7, simple task)

As shown in participant’s 7 comments, another typical tendency usually connected to state anxiety was to get fixated with something completely unrelated to the main task in which they were engaged, and instead, direct their attention to irrelevant thinking. Often, learners’ anxiety is expressed though worry about a present situation or task or a future potential outcome, and this worry affects their thought process.

Within this category, many respondents appeared to have experienced some
difficulty, and more specifically, regarding lexical knowledge, as they were not able to think and come up with the right words they were looking for.

“It made me very self-conscious and aware of how little applicable vocabulary I knew. I felt limited in my expression and was struggling to draw from the few words in my Spanish vocab, and it was frustrating because it seems like I could not remember a fraction of the words I should know.” (participant 7, complex task).

5.3.1.3 Thought and Idea Organization

Qualitative data reflected another cognitive function directly affected by state anxiety, which led to detrimental effects on participants’ performance, namely, ‘information processing’ or ‘idea organization’. Let us examine the following comment:

“This task left me flustered because I had to come up with a lot of vocabulary by myself. I can conjugate verbs in the past/present tense well, but it was difficult remembering verbs and nouns. I had to pause often between ideas. I was anxious because I couldn’t remember myself vocabulary words.” (participant 19, complex task)

As for participant 19, his state anxiety during the task often interrupted his normal thought process and the ideas that were coming to his mind.

Participant 3 (below) felt stress while narrating the simple story, and that seemed to have influenced the way he processed the information during the task by slowing down his train of thought.

“I struggled with explanation of the pictures because of the stress I felt. It was hard to process so much information quickly and in Spanish.” (participant 3, simple task)

5.3.1.4 Focus

Lastly, state anxiety in the current study also was reported to have a negative impact
on the participants’ focus and concentration during both the simple and the complex tasks. Participants’ 23, 44 and comments showed how stress made him/her lose concentration in while narrating the story.

“I felt stressed because I knew I was being recorded but I could not think of what to say, so that made me even more stressed and then made it even harder to know what to say. However, this task had more going on in it than the other storyline and that made coming up with what to say easier.” (participant 23, complex task)

“When I realized that I started the story wrong, I become flustered and it impeded my performance on some level because I was distracted by my mistake.” (participant 44, simple task).

“It decreased the quality of my performance because when I got flustered and confused, I'd lose my focus and couldn't remember how I planned to continue telling that story moving forward.” (participant 44, complex task)

The distribution of these processes across the two levels of task complexity was similar, and it was observed at equal levels between the simple and the complex task in the current study.

5.3.2 On the Three-Way Interface Language Anxiety — Cognition — Task Performance in Second Language Tasks

One of the main goals of the current study was to examine the interplay between the following participants’ dimensions: Language Anxiety — Cognition — Task Performance in L2 Spanish. This was done with the purpose of offering evidence of the dynamic interrelation between emotions, cognition and L2 performance by means of
manipulating cognitive load in pedagogic L2 tasks, as partially suggested by previous studies.

Critically, as revealed by the qualitative data, a strong dynamic complex relationship among the three components exists when cognitive complexity is manipulated in an instructional setting. The section below discusses the results found in relation to the three-way interface Language Anxiety — Cognition — Task Performance with illustrative examples extracted from participants’ comments.

5.3.2.1 Relationship ‘Language Anxiety — Cognition — Task Performance’ in L2 Spanish

The three-way interface between the state anxiety, the cognition and the L2 performance components was reflected in participants’ responses. As expected, results showed a clear relationship between state anxiety, cognitive processes and L2 performance (as perceived by participants) for the simple or the complex task. 37 participants (72%) made an explicit connection between language anxiety affecting their cognitive abilities, and consequently, their task performance. Among the most mentioned cognitive processes negatively impacted were participants’ thought process, memory, thought and idea organization, focus and concentration.

Language anxiety was perceived by participants as both a) the cause and b) the consequence of processing in the current study. A slowdown in the thought process, ability to focus, and interference with memory (particularly experiencing difficulty with lexical access), caused learners to feel state anxiety and apprehension during task performance. The following extract is a typical example of anxiety as the consequence of cognitive processes and poor L2 performance:
“I was definitely a bit flustered when I couldn't remember the words to say what I was thinking. I was less nervous speaking than in class, but usually in class I get instant feedback as to my errors or words I do not know so I can continue more easily, whereas in this situation I had no one to help me.” (participant 40, complex task)

In this case, feeling anxious comes as a result of not being able to remember the necessary words to complete the task successfully. However, let us consider the following comment:

“I felt anxious, flustered, and nervous when completing this task because I am not a strong Spanish speaker so anytime I have to speak in Spanish I get nervous and tend to forget the Spanish that I do know, and I get stuck when I do not know the correct vocab words to describe things. I also felt this way because there were other people in the room while I was speaking.” (participant 6, simple task)

In this case, the process was reversed, and language anxiety was experienced as the consequence of processing rather than the cause. Interestingly, data revealed these interesting patterns in these 72% participants.

Results further showed the cyclical relationship of Emotions - Cognition where a slowdown in the thought process provoked state anxiety (consequence), and in turn, this anxiety continued being the source of slower processing (cause). This two-way relationship often times incorporates the behavioral component in the equation as another variable that influences and is influenced by the other two components. The behavioral component refers to the participants’ task performance in Spanish as an L2.

In the following comment, the participant mentioned how his performance was impeded due to slow cognition, which in turn, was affected by state anxiety. The
relationship Language Anxiety — Cognition — L2 Task Performance is clearly reflected in the extract below:

“This affected my performance because it is harder for me to think and come up with the correct answers and sentences when I am feeling very anxious and I make more mistakes than I do when I am less anxious.” (participant 11, complex task)

Interestingly, when completing the complex task, participant 27 reflected the opposite direction in the relationship:

“Feeling frustrated led me to work a bit slower in telling the story. As I did not want to tell the wrong word, it took longer to come up with the right way to describe the picture”. (participant 27, complex task).

Feelings of state anxiety led the participant to process the information slower, and therefore, having a negative impact on language production.

The simultaneous interaction of these three components among them and their interrelation in a circular way was also evidenced in participants 23 and 25 when performing the cognitively complex task:

“I think that I was in a vicious circle of not knowing that to say then getting more stressed, so then I didn't know what to say even more and then my performance was affected very badly. (participant 23, complex task)

“I could not think of what to say, so that made me even more stressed and then made it even harder to know what to say. However, this task had more going on in it than the other storyline and that made coming up with what to say easier.” (participant 25, complex task)

Although these comments were only elicited by a few participants, and thus,
reflected in a small percentage of the data, it shows there is a trend that would be worth investigating in future studies more in depth.

5.3.3 On the Relationship between State Anxiety and Second Language Performance—as Perceived by Learners—When Cognitive Task Complexity is Manipulated (Research Question 4b)

To answer to RQ4b, results in the qualitative data demonstrated that, overall state anxiety had a negative effect during task performance (as perceived by participants) on the two levels of task complexity. All participants reported having at least one aspect or area of the task that induced language anxiety as concurrent to their linguistic performance. Some participants only reported anxiety in either the simple or the complex task, while others did it in both tasks.

In the majority of the cases, participants described their influence of language anxiety on their task performance as far below their linguistics capabilities in Spanish. Among the terms used to describe their performance were: ‘negative performance’, ‘incoherent’, ‘worse than expected’, ‘poor’, ‘terrible’, ‘awful’, ‘choppy speech’, among other terminology.

When examining each task separately, results revealed clear statistical differences between the simple and the complex task in the degree to which participants’ performance was considered to be affected in relation to state anxiety. Out of 51, 33 participants (64%)—above average—reported their task performance on the complex task was significantly poorer when compared to the simple task as a result—at least in part—to the language anxiety experienced. Below, an example is shown where a participant commented on his perception of his performance during the complex task.
“My stress and frustration seemed to be blocking out skills and words that I have learned before in class, which made my performance even worse because I stuck to using the same extremely limited set of words the whole time instead of elaborating the story line or having any creativity.” (participant 8, complex task)

For the remaining 34% of the participants, whereas 21% claimed their L2 performance improved for the complex task, for 13% of participants, overall, no statistical difference was observed between the simple and the complex task regarding their perceptions on their L2 performance. In other words, these participants reported that their performance was equally poorer as compared to their usual performance in Spanish.

5.3.3.1 Learners’ Perceptions on the Main Linguistic Areas Affected by State Anxiety

Qualitative data revealed participants’ task performance was affected by state anxiety and certain areas were more prominent than others. Results showed that the linguistic areas that were mentioned by participants with more frequency came down to three: a) ‘L2 fluency and productivity’, b) ‘L2 grammar and linguistic accuracy’, and c) ‘creativity’.

5.3.3.1.1 Debilitative and Facilitative Anxiety: Learners’ Perceptions on its Effects on Second Language Task Performance

For the debilitative type of anxiety—which led to poor task performance—results for the simple task indicated ‘L2 fluency and productivity’ as the most highly reported areas of poor task performance by 64% of the total number of participants; followed by ‘creativity’ 32% of the participants; and ‘L2 grammar and linguistic accuracy’, mentioned by 20% of participants.

For the complex task, ‘L2 fluency and productivity’ were present in a higher
percentage of the participants’ responses (72%) as compared to the simple task. ‘L2 grammar and linguistic accuracy’ came up as the second most important area by state anxiety (27%). Lastly, ‘creativity’ was only mentioned by 2% percent of the total number of participants.

The following example has been taken from participant’s 24 responses. Let us consider his perceptions on how state anxiety affected his task performance:

“It made me stumble on certain pictures and take long pauses when I could not remember the word and then I was afraid to be more creative because I did not want to mess up more of the words or forget the vocabulary”. (participant 24, simple task)

As can be seen, participant 24 referred to his speech as being affected by long pauses. Additionally, the participant also highlighted his lack of creativity as a major problem of his oral performance. Below, two examples shows learners’ perspective on how their performance was negatively impacted in relation to L2 grammar and linguistic accuracy.

“I definitely stumbled over words and my conjugation of verbs was way off.” (participant 10, simple task)

“I wasn’t able to process what I wanted to say as quickly, and it was harder to decide what to say and conjugate verbs.” (participant 32, simple task)

As reflected in participants 10 and 32, verb conjugations were one of the most highly frequently mentioned problematic areas with respect to L2 grammar. Overall, participants perceived their performance was negatively affected in the same linguistic areas. However, for the complex task, ‘L2 fluency and productivity’ and ‘creativity’ obtained higher percentages (72% and 32%), over the simple task, (64% and 2%,
respectively).

When comparing participants’ comments between the cognitively simple and the cognitively complex task, results further showed an important trend in relation to participants’ facilitative anxiety. Overall, from those participants who concluded their performance was overall good, differences between the two tasks were observed regarding the linguistic areas most frequently mentioned.

For the complex task, 85% of participants indicated their task performance as significantly more fluent, whereas for the simple task, only 29% of participants reported their speech benefitted in L2 fluency and the amount of language produced (productivity).

In line with previous results, overall, the complex task led the majority of participants to perceive their performance to be far more creative due to the higher number of elements present in the storyline. The following extract presents an example of a participant’s comment, in which the idea of room for creativity and a positive effect on L2 fluency was present.

“I think I talked more smoothly, still with pauses and unsure moments but there were less complete silences. I also did not have as much time to think about all of my verb conjugations because there was so much going on in each picture for my brain to process. I also felt more comfortable with the researcher listening so even when I realized I was using the wrong vocabulary I did not stumble and fixate on it but instead kept talking because there was a lot to talk about.” (participant 24, complex task).

Despite the differences observed between the two levels of task complexity, overall, results demonstrated the vast majority of participants seemed to have experienced a higher degree of debilitative state anxiety (over the facilitative type), on L2 fluency and
productivity, L2 grammar and linguistic accuracy, and creativity.

5.3.3.2 Language Anxiety Coping Strategies

‘Coping mechanisms and strategies’ were revealed as a common theme present in participants’ comments to deal with language anxiety and achieve a good performance in both narrative tasks.

Interestingly, results revealed similar levels of coping anxiety skills between the two levels of task complexity. For the simple task, 39% of the total number of participants and 41% for the complex task reported on the use of coping skills. Therefore, although for the cognitively complex task, a slight increase in the number of participants was evidenced over the cognitively simple task (in relation to coping skills), this difference did not seem to be statistically significant.

In what follows, the most prominent language anxiety coping skills will be described with illustrative examples taken from the data. According to the analyses, the most relevant category that emerged in the data were ‘avoidance strategies’ (group 1) mentioned by 29% of participants, followed by ‘other emotional compensatory strategies’ (group 2) reported by 11%.

5.3.3.2.1 Group 1: ‘Avoidance Strategies’

‘Avoidance strategies’: (Pappamihiel, 2002) were employed and reported by a number of participants at a lower degree. These refer to a response that participants, experiencing anxiety and fear of making errors in their speech, use to avoid taking risks in their oral performance (in this case). Let us consider the following examples:

1. “I believe using basic grammar helped me feel less anxious because I feel comfortable using these tenses”, (participant 12, simple task);
2. “I resorted to much simpler structures, although I know how to use different and more complicated structures”, (participant 13, simple task)

3. “Because I was so anxious about grammar and vocabulary, I basically described what was going on instead of being creative” (participant 15, complex task);

4. “I realized I told the story very quickly because I didn’t want to mess up more” (participant 16, simple task”)

5. “I used the present tense because that is the tense which I am most confident speaking in. I also tried to keep the storyline simple and linear, so I did not run into overly complicated grammar when I was speaking. I also attempted to use words I knew to describe the story.” (participant 6, complex task)

As observed in the first two extracts, state anxiety was shown in participant 12 and 23’s resorting to use basic structures in Spanish to avoid making errors in the language. Out of fear, they decided not to take linguistic risks by employing language structures they felt comfortable with, since that helped them feel more confident and deal with their state anxiety while performing. Likewise, participants 15 (in 3), participant 16 (in 4) and 6 (in 5) failed to be creative and spoke much faster than usual, respectively, because they were worried about making grammatical mistakes.

5.3.3.2 Group 2: ‘Other Emotional Compensatory Strategies’

a. ‘Alternative linguistic resources’: As part of these strategies, to help participants cope with anxiety, results showed that occasionally participants tried to develop their pragmatic competence, by applying mechanisms—such as circumlocution, paraphrasing, and searching for alternative ways—in order to communicate what was happening in the story when the necessary vocabulary and expressions were missing:
6. “searching for alternative ways to tell the story made me take significantly longer because I would have to figure out different vocab tools to say words I didn’t know” (participant 24, simple task);

b. ‘Self-doubt’ & ‘Low self-esteem’

Data revealed other strategies participants employed to cope with their language anxiety during performance were ‘self-doubting’, and ‘low self-esteem’. These were mentioned by some of the participants as psychological tendencies related to anxiety that led to a decline in learners’ task performance:

7. “Because I feel that I am not good at speaking in Spanish, this task was stressful. During the task, I was trying to perform as best as I could. This made me anxious because I knew that since I was not very good at speaking, what I was saying might not make much sense, or might seem like I know very little about speaking Spanish.” (participant 29, complex task)

8. “I felt quite stressed while completing this activity and it was surprising for me because I didn't think I would feel this way after I was given instructions. I felt very nervous throughout the whole exercise and I was constantly doubting myself and my performance. I think I felt so nervous because I didn't feel that I knew of enough vocabulary to tell the story fully. An interesting thing to note that I first felt confident with this exercise because the information prior to starting the exercise made me realize that I have been learning Spanish for a number of years. However, the actually doing the task made me feel as if I wasn't as good at Spanish as I had originally thought.” (participant 41, complex task)

c. ‘Self-motivation, positive thinking to over compensate anxiety’

Other conscious psychological coping strategies to reduce anxiety and
overcompensate the feelings were observed in participants 15, 17 and 23 during the simple and the complex task:

10. “I tried not to stress so much this time and think in Spanish” (participant 15, simple task)

11. “It was only when I really relaxed did the process became more fun” (participant 23, complex task);

12. “When I am anxious it is harder to think, so when I am speaking I have to overcome that hurdle and relax into the task” (participant 17, simple task).

5.3.4 Summary of Results for Research Question 4

RQ4 addressed the relationship between state anxiety, cognitive processes and L2 performance as perceived by participants. Overall, results from RQ4 in the present study revealed that state anxiety, fundamentally had a detrimental on cognition, and subsequently, on L2 performance as experienced by L2 learners.

All participants mentioned at least one aspect of the task that they considered was affected by their levels of anxiety experienced. When looking at the differences between the two tasks’ complexity levels, it was found that a majority of participants (64%) reported a negative impact of language anxiety (debilitative anxiety) in their performance on the complex task. On the other hand, 21% of participants claimed a positive effect of anxiety in their L2 performance (facilitative). These participants argued that although some anxiety was triggered (especially due to challenging vocabulary and grammatical constructions elicited in the tasks), overall their performance was better when compared to the simple task.
Lastly, a smaller percentage (13% of participants) did not report meaningful differences between the cognitively simple and the cognitively complex tasks in their task performance. For this group, participants did not feel their performance suffered to a higher degree in the complex task over the simple.

72% of the qualitative data revealed a strong connection in the Language Anxiety — Cognition — L2 Performance. This relationship was confirmed as equally relevant in both levels of task complexity. Importantly, a pattern was discovered in relation to the cognitive functions in which state anxiety appeared to have had a more prominent effect. Overall, participants reported experiencing difficulties mostly in: remembering the necessary vocabulary required to tell the stories (short-term memory), slowing down the thought process, information processing, and breaking their concentration from the main task with irrelevant unrelated thinking. Overall, no statistically significant differences were found between the two tasks regarding cognitive processes.

Among the most relevant results, the linguistic areas revealed as most negatively affected overall were: L2 fluency and productivity—mentioned by 64% of the total number of participants in the simple task, and by 72% in the complex task; followed by grammar—reported by 20% of participants in the simple and 27% of participants in the complex task; and lastly, creativity—found in 32% of participants’ comments in the simple task, and only in 2% in the complex task.

Overall, learners considered that for both the simple and the complex task, their narratives were considerably shorter, with a much higher amount of hesitations and more grammatical errors, and lastly, less creative due to the state anxiety experienced during task performance.
To conclude, results revealed participants used coping anxiety strategies in a similar degree in both tasks. In this respect, avoidance strategies were pointed out by almost a third (29%) of the participants.

5.4 Results for Second Language Performance: Complexity, Accuracy and Fluency Measures across the Two Cognitive Task Complexity Levels

RQ5 sought to examine how presumed cognitive task complexity was related to L2 performance on L2 Spanish learners in order to explore the interplay between Language anxiety — Cognition — L2 Performance. To assess task performance, the outcome measures for fluency, accuracy and syntactic complexity were compared for the two levels of cognitive complexity. Scores for each of the performance measure were explored for means, SDs, confidence intervals, effect sizes and inferential statistical procedures to examine possible differences between the two task complexity conditions. In the following section, linguistic trends are reported for mean length of AS-unit, mean length of clause, number of clauses per AS-unit, weighted clause ratio, total number of words, total number of syllables (excluding repetitions, repairs, and false starts), and finally, pruned speech rate measured by (a) words per minute and (b) syllables per second.

5.4.1 Syntactic Complexity Measures

Before calculating frequency statistics, data was inspected for missing cases, outliers and normality of distributions. Out of the 51, 1 participant was identified as missing, leaving 50 participants ($N = 50$). With regard to the units of analyses, one extreme bivariate outlier and two univariate outliers were found for clauses, and four univariate outliers for AS-units. In relation to the three syntactic complexity measures, one univariate extreme outlier was identified for mean length of As-units (MLASU), three univariate (one
extreme) outliers, and one bivariate for mean length of clause (MLC), and two univariate extreme outliers for number of clauses per AS-unit (C/AS-units). After removing outliers, it was concluded that normality of distributions was met for all syntactic complexity measures; and the total number of participants for each measure was: \( N = 47 \) for clauses and \( N = 46 \) for As-Units, \( N = 49 \) for mean length of AS-units, \( N = 46 \) for mean length of clause, and \( N = 48 \) for clauses per AS-units.

Table 20 presents the results of the descriptive statistics on the total number of clauses and AS-units produced by participants for the two tasks’ complexity cognitive conditions and the three syntactic complexity measures. Confidence interval error bars in Figures 26, 27, and 28 showed the same information visually. For the three outcome measures, meaningful differences were observed between the task imposing lower cognitive demands as compared to the task with higher cognitive demands. Meaningful differences were revealed in relation to global complexity (mean length of AS-units); followed by amount of subordination (number of clauses per AS-units); and finally, phrasal complexity (mean length of clause) between the between the simple and the complex task. For the three measures of syntactic complexity, the mean values for the complex task were higher than the upper 95% confidence intervals, demonstrating a statistically significant difference between the two cognitive complexity levels. In other words, the complex task led participants to produce significantly more complex language syntactically.
**Table 20:** Descriptive statistics for syntactic complexity for the simple and the complex tasks.

<table>
<thead>
<tr>
<th>Task</th>
<th>Nº Clauses</th>
<th>AS-Units</th>
<th>MLASU</th>
<th>MLC</th>
<th>C/AS-Units</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>C</td>
<td>S</td>
<td>C</td>
<td>S</td>
</tr>
<tr>
<td>N</td>
<td>47</td>
<td>46</td>
<td>49</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>Mean</td>
<td>13.08</td>
<td>16.97</td>
<td>11.80</td>
<td>12.63</td>
<td>5.99</td>
</tr>
<tr>
<td>SD</td>
<td>3.96</td>
<td>6.11</td>
<td>2.75</td>
<td>4.20</td>
<td>2.00</td>
</tr>
<tr>
<td>Min</td>
<td>4.00</td>
<td>7.00</td>
<td>6.00</td>
<td>5.00</td>
<td>2.20</td>
</tr>
<tr>
<td>Max</td>
<td>23.00</td>
<td>32.00</td>
<td>18.00</td>
<td>23.00</td>
<td>11.00</td>
</tr>
</tbody>
</table>

**Figure 26:** Confidence intervals for the global syntactic complexity measure—mean length of AS-units (MLASU) — for the simple and the complex task.
Figure 27: Confidence intervals for the phrasal syntactic complexity measure—mean length of clause (MLC) — for the simple and the complex task.
A one-way repeated measures ANOVA was conducted individually on each complexity measure and the same results were confirmed. Meaningful (statistically significant) differences between the simple and the complex task were revealed for global syntactic complexity (MLASU), \( (N = 49) \, F(1,48) = 74.32, p = 0.000, \) partial Eta squared = 0.60, for clausal complexity (MLC) \( (N = 46) \, F(1,45) = 36.53, p = 0.000, \) partial Eta squared = 0.44, and amount of subordination measure (C/AS-units) \( (N = 48) \, F(1,47) = 56.26, p = 0.000, \) partial Eta squared = 0.54. Cohen’s d (1988) revealed large effect sizes for global syntactic complexity \( (d = 1.29) \), for subordination \( (d = 1.20) \), and for phrasal complexity \( (d = 0.9) \).
In sum, the results indicated that participants’ stories in the complex task contained structurally more sophisticated and complex language in terms of the length of their syntactic constructions (MLASU), subordination (C/AS-units), and a higher number of nominalizations (such as noun phrases) (MLC).

5.4.2 Accuracy Measure

A similar pattern was observed when examining the relationship between task complexity and accuracy. For weighted clause ratio, out of 51 participants, data from 1 missing case was lost. Participant 15 was removed from the data set as a univariate outlier for the accuracy measure. After the outlier was removed, as seen in Figures 29 and 30, normal distributions were met for the WCR on the two levels of cognitive task complexity. For the simple task, only a minimal negative skewness was found, which indicated most participants’ scores were slightly placed towards the left of the mean. No severe skewness or kurtosis was found on the score distributions for WCR on the complex task.
**Figure 29:** Histogram for the accuracy measure—*weighted clause ratio* (WCR)—on the cognitively simple task.

**Figure 30:** Histogram for the accuracy measure—*weighted clause ratio* (WCR)—for the cognitively complex task.
Table 21 shows the mean number of clauses per error type, calculations for the weighting on each error level, raw total scores, and weighted clause ratio for the cognitively simple and the cognitively complex task.

**Table 21:** Mean number of clauses per error levels, weighting error clauses, raw total score, and *weighted clause ratio* (WCR) for the two levels of cognitive task complexity.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Simple Task</th>
<th>Complex Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of clauses</td>
<td>13.94</td>
<td>18.72</td>
</tr>
<tr>
<td>Entirely accurate clauses</td>
<td>3.84</td>
<td>3.35</td>
</tr>
<tr>
<td>Clauses at Level 1</td>
<td>3.92</td>
<td>4.35</td>
</tr>
<tr>
<td>Clauses at Level 2</td>
<td>3.38</td>
<td>7.00</td>
</tr>
<tr>
<td>Clauses at Level 3</td>
<td>2.65</td>
<td>4.00</td>
</tr>
<tr>
<td>Weighting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entirely accurate clauses × 1.00</td>
<td>3.84</td>
<td>3.35</td>
</tr>
<tr>
<td>Level 1 × 0.80</td>
<td>3.13</td>
<td>3.48</td>
</tr>
<tr>
<td>Level 2 × 0.50</td>
<td>1.69</td>
<td>3.50</td>
</tr>
<tr>
<td>Level 3 × 0.10</td>
<td>0.33</td>
<td>0.40</td>
</tr>
<tr>
<td>Raw total</td>
<td>8.99</td>
<td>10.72</td>
</tr>
<tr>
<td><strong>Weighted Clause Ratio</strong></td>
<td>0.64</td>
<td>0.55</td>
</tr>
<tr>
<td><em>(raw total divided by total clauses)</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As displayed in Table 22, information on the means, *SDs*, minimum and maximum
score values are presented for the two cognitive complexity levels in relation to the type of errors produced and their gravity. Descriptive statistics revealed a very different pattern for the accuracy measure. Over a range of 0.24 (minimum score) to 1 (maximum score) for the simple task, and 0.25 (minimum) to 0.85 (maximum) for the complex task, mean scores for the simple task were situated at $M = 0.64$, significantly higher than mean scores for the complex narrative, situated at $M = 0.55$.

Statistical significant differences between the two task complexity conditions were observed for the WCR ratio. However, this time, participants significantly improved their performance in the simple task, which showed the least number of errors. On the contrary, increasing the cognitive demands of the task led to a higher number of errors, and the difference was meaningful.

**Table 22**: Descriptive statistics for the accuracy measure—*weighted clause ratio* (WCR)—based on the simple and the complex task.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Simple Task</th>
<th>Complex Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Mean</td>
<td>0.64</td>
<td>0.55</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>Min</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>Max</td>
<td>1.00</td>
<td>0.85</td>
</tr>
</tbody>
</table>

As shown in Figure 31, confidence intervals confirmed the meaningful differences between the two levels of cognitive complexity for the accuracy measure. The upper bound CI of the simple task exceeded the mean scores for the complex task indicating that participants were more accurate in the simple task.
Figure 31: Error bars for the accuracy measure—*weighted clause ratio* (WCR)—based on the simple and the complex task.

One-way Repeated Measures ANOVA was run for the accuracy measure based on the two levels of task complexity, and statistically significant differences were revealed ($N = 49$) $F(1,48) = 27.47$, $p = 0.000$, partial Eta squared = 0.36. Cohen’s d (1988) revealed a large effect size for task complexity ($d = 0.65$) on accuracy.

5.4.3 Fluency Measures

Before conducting analyses, descriptive statistics for the four fluency measures were calculated. For the productivity measures, one participant’s score was missing,
leaving a total of 50 participants \((N = 50)\). Three univariate outliers were found for total number of words (TNW), and two univariate outliers for the total number of syllables (TNS). For the pruned speech rate, no severe outliers were revealed for words per minute (W/min), and two outliers (one bivariate and one bivariate) were identified for syllables per second (Syl/sec).

After outliers were withdrawn from the data set, for the two productivity measures, histograms and descriptive statistics revealed normal distributions for both the simple or the complex tasks. For total number of words (TNW) and (TNS), Figures 32, 33, 34 and 35 showed distribution of scores for the simple and the complex task, respectively.

![Histogram for total number of words (TNW) for the simple task.](image)

**Figure 32:** Histogram for total number of words (TNW) for the simple task.
Figure 33: Histogram for total number of words (TNW) for the complex task.

Figure 34: Histogram for total number of syllables (TNS) for the simple task.
Figure 35: Histogram for total number of syllables (TNS) for the complex task.

For TNW, scores ranged between a minimum score of 29.00 and a maximum of 142.00, with a mean score of 84.42 for the simple task; and a minimum value of 59.00, a maximum of 241.00, and a mean of 138.08 for the complex task. For TNS, scores ranged from 47.00 to 262.00, with a mean score of 139.00 for the simple story. For the complex storyline task, the minimum score was of 93.00, the maximum of 417.00, and the mean score of 231.16 for the complex task. Means and SDs (see —3), and non-overlapping confidence intervals (see Figures 36 & 37) demonstrated statistically significant differences between the cognitively simple and the cognitively complex task for the two productivity measures, TNW and TNS.
Table 23: Descriptive statistics for the productivity measures for the simple and the complex task.

<table>
<thead>
<tr>
<th></th>
<th>Total number of words (TNW)</th>
<th>Total number of syllables (TNS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Complex</td>
</tr>
<tr>
<td>N</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>Mean</td>
<td>84.42</td>
<td>138.08</td>
</tr>
<tr>
<td>SD</td>
<td>37.29</td>
<td>47.81</td>
</tr>
<tr>
<td>Min</td>
<td>29.00</td>
<td>59.00</td>
</tr>
<tr>
<td>Max</td>
<td>142.00</td>
<td>241.00</td>
</tr>
</tbody>
</table>

Figure 36: Confidence intervals for total number of words (TNW) based on the two complexity levels.
As can be seen in Figures 38, 39, 40 and 41 for pruned speech rate, normal distributions were also demonstrated for the speech rate measures after removing the outliers. Participants’ scores ($N = 48$) for the simple and the complex task showed similar trends to the productivity measures. Histograms and descriptive statistics indicated normal distributions, with very slight levels of skewness and kurtosis, for the two speech rate measures.
Figure 38: Histogram for words per minute (W/min) for the simple task.

Figure 39: Histogram for words per minute (W/min) for the complex task.
Figure 40: Histogram for syllables per second (Syl./sec) for the simple task.

Figure 41: Histogram for syllables per second (Syl./sec) for the complex task.

From Table 24, it can be concluded that the cognitively complex task yielded a statistically significant higher number of words per minute $(M = 37.09)$, and syllables per
second \((M = 0.96)\) as compared to the simple task, \((M = 44.54)\), and \((M = 1.21)\) for words per minute, and syllables per second, respectively. As shown by the error bars (see Figures 42 & 43), the no overlapping between the confidence intervals for the three rate measures indicated linguistic meaningful differences regarding their L2 fluency between the cognitively simple and the cognitively complex task.

**Table 24:** Descriptive statistics for the speech rate measures for the simple and the complex task.

<table>
<thead>
<tr>
<th></th>
<th>W/min</th>
<th>Syl/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Complex</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>37.09</td>
<td>44.54</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>15.90</td>
<td>15.00</td>
</tr>
<tr>
<td><strong>Min</strong></td>
<td>8.50</td>
<td>16.00</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>83.33</td>
<td>82.05</td>
</tr>
</tbody>
</table>
Figure 42: Confidence intervals for words per minute (W/min) for the simple and the complex task.
The results of a one-way repeated measures ANOVA revealed that these differences were indeed statistical for total number of words (N = 48) $F(1,47) = 89.82, p = 0.000$, partial Eta squared = 0.65, total number of syllables (N = 48) $F(1,47) = 130.85, p = 0.000$, partial Eta squared = 0.73, pruned rate of speech for words per minute (N = 48) $F(1,47) = 30.80, p = 0.000$, partial Eta squared = 0.39, and syllables per second (N = 48) $F(1,47) = 47.58, p = 0.000$, partial Eta squared = 0.50. Cohen’s d (1988) revealed large effect sizes for total number of words ($d = 1.41$), for total number of syllables ($d = 1.82$), for words per minute ($d = 0.81$), and for syllables per second ($d = 1.04$).
5.4.4 Summary of Results for Research Question 5

RQ5 investigated the relationship between cognitive task complexity and L2 performance, as assessed by CAF measures. Overall, results indicated that participants benefited from the cognitively complex task in terms of fluency and syntactic complexity. For accuracy, the trade-off effect was revealed between syntactic complexity and accuracy.

Specifically, results demonstrated that for syntactic complexity the intended-to-be complex task promoted significantly higher levels of global complexity, subordination, and clausal complexity. Similarly, for the fluency measures, results revealed the cognitively complex task promoted both productivity and speech rate. For the productivity (amount of language produced) measures, a statistically significant increase was shown for the total number of words and syllables produced. Likewise, for pruned speech rate, the number of words per minutes and syllables per second were statistically significantly higher as compared to the simple task. Finally, for accuracy, the opposite pattern was observed. Participants produced a meaningful decrease in the amount of accurate language produced in the designed-to-be complex task over the designed-to-be simple task. Overall, an increase in the cognitive demands of the task also led to a much higher number of errors.
CHAPTER 6: DISCUSSION

In this chapter, I summarize the results presented in chapter 5, and these are discussed in relation to each research question. Important patterns are highlighted to shed light on our understanding of the construct of cognitive task complexity, foreign language anxiety and their joint effect on learners’ task performance in the L2. Evidence from the results presented above contributes to cognitive psychology literature and language anxiety research in both the SLA and TBLT strands of research. Supporting the recent turn on language anxiety research, findings of the current dissertation provide further evidence for the complex and dynamic nature of language anxiety (as a learner’s individual factor), its dynamicity and fluctuation over time, and how its behavior is manifested during L2 task performance. The study makes a further step in establishing the connection between this affective factor, cognitive processes and the quality and quantity of linguistic performance from the learner’s subjective point of view in L2 tasks.

6.1 Research Question 1. Validating the Intended Relationship between Task Complexity and Cognitive Demands Imposed by the Two Tasks

The first RQ (RQ1) had a two-fold purpose. Firstly, it aimed to validate the actual cognitive complexity (versus the assumed complexity) of the tasks to ensure the two levels of complexity indeed posed distinct cognitive task demands. This allowed to a) address the potential relationship between state anxiety and cognitive task complexity, and b) partially examine the interplay Language Anxiety — Cognitive Processes — L2 Performance within the participants. Second, it sought to contribute to the recent task complexity literature that attempted to validate the cognitive demands of tasks by introducing independent cognitive load methodologies from cognitive psychology (Sweller, 1994) into the TBLT strand of
research. As explained in the literature review section, a significant number of SLA researchers (e.g. Norris & Ortega; 2003; Norris, 2003; and Sasayama, 2016) recommended the use of a variety of independent cognitive load methodologies to validate the assumed cognitive demands of L2 tasks.

Before situating the current findings in relation to previous research on task complexity, it is important to remind the reader that the cognitively simple and the cognitively complex tasks employed in the current study were originally adapted from Hill (1960) (“the alarm clock”) and Elder and Iwashita (2005) (“the dog”), respectively; and later employed by Sasayama’s (2015, 2016) operationalization of cognitive task complexity. As explained in the research design chapter, the selection of the simple and the complex task in the current study was based on Sasayama’s (2015) dissertation’s findings (2015) and her 2016 study on validating task complexity. In her study, the researcher employed a variety of cognitive load measures—task complexity, mental effort, time estimation and dual-task methodology—on L1 and L2 learners of Japanese to validate the cognitive demands of tasks. It is important to note that, for the L2 learners, she found large effect sizes (Cohen, 1988), but only for the subjective measures, task difficulty (d = 0.83) and mental effort (d = 1.05) in the low proficiency group.

A different pattern was observed for dual-task methodology and time estimation, where although statistically significant differences were found, the effect sizes observed for these two measures were medium to small, $d = 0.34, d = -0.39$, respectively. The effect sizes served as the main rationale for choosing the subjective measures in the current dissertation for validation of the cognitive demands.
Overall, findings revealed a complicated relationship between the two levels of cognitive task complexity and the cognitive load measures employed in the current study. First, contrary to expectation, quantitative and qualitative results showed a lack of relationship between the two levels of cognitive complexity on the task difficulty measure. Participants’ self-assessed ratings on task difficulty did not seem to vary significantly across the two tasks. In the same line, qualitative results showed that although a large number of participants reported the complex task to be significantly more difficult for them, surprisingly, a very similar number of participants (the other half) mentioned in their comments that the simple task had been perceived as significantly more difficult. This pattern is partially consistent with Sasayama (2015, 2016). Using the same tasks in her dissertation, she found both tasks being reported as difficult and requiring more mental effort for a large portion of her participants. In line with these results, in the current study, many participants commented that for the simple story, it had been extremely hard to be creative and come up with good ideas for the plot, which made the task more difficult. Interestingly, these participants attributed this task difficulty to the little action and the few details contained in the story. The current finding seems to go in line also with Kim et al. (2015). In their study, the researchers manipulated task complexity by increasing the reasoning demands of the task and found the two tasks were rated as equally difficult.

These findings strongly suggest that the relationship between task difficulty and task complexity will be dependent on the type of task as well as the particular operationalization employed by the researcher. In this respect, results in the current study indicate that at least for certain types of tasks—picture description tasks—an increase of
the conceptual input and the linguistic demands does not necessarily translate to a more difficult task.

For time estimation, results indicated no statistical difference between the two levels of cognitive task complexity in relation to the learners’ perception of time spent on the two tasks. Specifically, the simple task showed the longest negative difference \( M = -13.93 \), in relation to an increase in the mean scores observed for the cognitively complex task \( M = -7.44 \) when comparing the two levels of task complexity. Importantly, the relationship showed an increase in the mean scores observed for the cognitively complex task in relation to the simple task. Hence, the estimated time on task was perceived to be shorter on both levels of task complexity than their actual time; however, surprisingly, the more complex showed the smallest difference and this different was insignificant statistically. These findings appear to be only partially consistent with previous task complexity studies (Baralt, 2010, 2013; Sasayama, 2016) that employed time estimation as a measure of cognitive load.

In Sasayama’s study, the difference between the estimated time and the actual time on task was also negative for the simplest and the complex levels, suggesting that participants perceived the time spent on task to be significantly shorter than their actual time on task. However, this relationship was the opposite when comparing the two levels of task complexity. In particular, for the low level of proficiency, whereas the simple task (Task 1) yielded the second shortest negative difference between the actual time of task and the judged time \( M = -23.18 \), the cognitively complex task yielded the longest negative difference between participants’ estimated time and their actual time on task \( M = -39.80 \). The author concluded that these findings were in line with the prospective approach to time estimation.
estimation according to Block et al. (2010), in which the estimated time tends to decrease as cognitive complexity increases. As highlighted in the literature review section, this proposal has been postulated in connection to the attentional-gate model (Zakay & Block, 1997). The time estimation is perceived to be shorter as task complexity increases since, according to the dual-task methodology, there are fewer attentional resources available as two tasks are being performed simultaneously. Contrary to these findings, although the current findings revealed time estimated was shorter when compared to actual time on task, increasing task demands did not seem to lead to perceived shorter time. In this case, the time judged decreased for the simple task.

Overall, task complexity led to an increase in time estimation, but this increase was not meaningful, and thus, the prospective approach was not supported in this case. Importantly, the current pattern also contradicted the retrospective approach observed in Baralt (2010, 2013). In this case, the researcher found that the estimated time on the most complex task was significantly longer than learners’ actual time on the task, indicating a positive difference between time estimation and actual time on task. It is important to remember that the retrospective approach predicts the opposite relationship between length of estimated time and cognitive load (Block et al. 2010). In this case, this relationship is accounted by the contextual change hypothesis (Zakay & Block, 1997). This model relies on memory (rather than attention) to explain the relationship between the perception of time and the actual time on task, hypothesizing that retrospective judgments of time involve information retrieval stored in memory about changes that occurred during task performance. Also contrary to Baralt’s findings, the current pattern does not appear support the retrospective approach either.
The current dissertation partially agrees with Sasayama (2015, 2016), nonetheless, in that time estimation does not seem to be as responsive and effective to distinguish between the cognitive load imposed by L2 tasks on lower levels of proficiency as it is for higher levels of proficiency. On this note, it is important to remember that the researcher found a medium to small effect size \( d = -0.39 \) for this measure when compared Task 1 and Task 4 in the lower level of proficiency.

Overall, the current findings are in accordance with Sasayama (2015, 2016) in all measures of cognitive load where differences between the simple and the complex task are less consistent in distinguishing among low proficiency learners than it is among high levels of proficiency. The author observed that as the level of L2 proficiency decreased, the patterns and differences observed in the cognitive load imposed by the distinct levels were not as important. In this sense, in the current dissertation, this finding is also supported, as learners also were novice learners (as determined by their institutional proficiency status, enrolled in their third semester of Spanish). The low level of proficiency and the task appear to be important factors determining the insignificant differences between the complexity levels.

Despite no statistical significant differences found for two (out of three) of the measures employed, namely, task difficulty and time estimation, findings seem to indicate that the complex task indeed seemed to have imposed a statistically significantly greater cognitive load at least for one of the three independent methodologies employed—mental effort—as compared to the simple task. This pattern supports Révész et al. (2016) in their findings, where statistical differences were only found on the decision-making and map tasks and not on the narrative picture-description task.
As highlighted above, for comparability reasons, the decision of using ‘number of elements present in the task (+/- few elements)’ was based on the fact that it has been Robinson’s (2001, 2007) most widely investigated resource-directing variable in the manipulation of conceptual and cognitive demands.

Consistent with Sasayama (2016), results of the current dissertation demonstrated that increasing the ‘number of elements’ (characters) contributed in a meaningful manner to impose significantly higher cognitive demands on L2 learners. Therefore, among the most important findings, the present study suggests that increasing the number of elements in a picture-description task affects task complexity. These results appeared to support Robinson’s (2001a) prediction—at least partially—on the relationship between conceptual information and cognitive demands of L2 tasks.

Qualitative results provided further evidence on the more strongly associated sources to task difficulty, namely, in order of importance: linguistic demands (for 78% of participants) and conceptual information (number of characters) (for 58% of participants). These findings are in accordance with Sasayama (2015), which revealed conceptual input and code complexity as the most prominent superordinate categories for task difficulty. Particularly, they appear to support the claim that as proficiency level increases, linguistic demands do not play the same important role as they do for the more novice learners.

In relation to the linguistic demands, vocabulary was identified as the most important factor affecting the perceived easiness or the difficulty of the task, also in line with recent studies on task complexity (Gurzynski-Weiss & Révész, 2016; Sasayama, 2016).
Results further demonstrated that ‘L2 grammar’, and more specifically, verb tenses and conjugations, were mentioned as the second most reported category by 21% of participants, which is not surprising considering participants’ (low) proficiency level. Aside from linguistic demands, it is interesting to note that there were other factors (as expressed by 51% of participants) that seem to have influenced participants’ perceptions in significant way of how difficult the task was experienced. Within these factors, task mode (oral versus written), the absence of planning time, and the type of task were among the most problematic issues mentioned by participants. This pattern may suggest that for lower proficiency levels, providing learners with enough preparation time prior to task performance, especially in oral production tasks, is of utmost important, as this can have a considerable impact on learners’ perceptions of the difficulty of the task in hand.

As recommended by recent research (Gurzynski-Weiss & Révész, 2016), the current study further addressed the need to examine perceptions on task difficulty from the learners’ point of view (POV). The patterns observed in this dissertation seem to be very consistent with Gurzynski-Weiss and Révész (2016) in relation to the task factors that were proposed as the major contributors to task difficulty. The researchers explored the most important sources that, from the teachers’ perspective, affected task difficulty. The study demonstrated that linguistic demands, followed by conceptual information, were by far the most important factors that seemed to have a major impact on task difficulty.

In light of the current findings, the efficiency of task difficulty and time estimation to assess cognitive load appears to be questionable within lower levels of proficiency and particularly for this type of task. This may be due to the fact that there may not be a straightforward relationship between the difficulty of a task and the amount of cognitive
load imposed. As highlighted by almost half of the participants and also observed in Sasayama (2016), increasing the number of elements present as part of the conceptual information of the task appears to be beneficial for low proficiency learners. This may be explained because higher cognitive load does not necessarily translate to more difficulty, especially when the task provides learners with a particular level of creativity. It is clear that the monologic condition of the task (Robinson, 2005), where no interlocutor is involved, provides learners with many fewer opportunities to be creative. Manipulating the conceptual input may give learners more chances to focus on alternative linguistic structures when they are lacking the adequate linguistic resources to describe certain challenging information and communicate in the L2.

The type of task, as a narrative picture description task, seems to also influence the perception of difficulty. Narrating a story without preparation and without a list of vocabulary to aid learners’ performance was an observation made by many participants and considered as a major source of difficulty.

Regarding mental effort, results in the current dissertation seem to indicate a more linear relationship between the amount of brain power exerted in the task and the cognitive demands of the task. Findings suggest that participant’s perceived mental effort for this particular operationalization was strongly related to the level of cognitive task complexity. Approximately 60% of the participants reported that the cognitively complex task required more mental effort over the simple task. For the majority of these participants, the complexity was attributed to a) linguistic demands (63% of cases), and b) a higher number of characters depicted in the pictures and “more stuff going on” (32% of cases) in the
complex task that seems to have forced them to put much more mental effort in their L2 performance.

The current study also suggests that both Robinson’s Cognition Hypothesis and Skehan’s Limited Attention Capacity model provide a partial view of the role that the various components that intervene in task design and L2 learners’ processes are depicted. To explain, although both proposals provide huge contributions in explaining the relationship between cognitive task demands and L2 task performance, they also show some limitations. For example, Robinson’s triadic model fails to address the linguistic demands of L2 tasks (represented in Skehan as ‘code complexity’). Conversely, Skehan’s model addresses the complexity of the language elicited in the task by incorporating the ‘code complexity’ component but fails to include learners’ individual factors, represented in Robinson’s CH.

In connection to the cognitive demands imposed by the tasks and the amount of conceptual information manipulated by the researcher, results revealed for certain learners having more characters and more action depicted in the story was experienced as especially more challenging to process than for others. This may be related to differences in learners’ short-term memory, which may lead to learners processing the information differently. Working memory capacity (reference), as well as other learners’ factors, results in the use of different strategies to process the information as the same time as performing the task. It was further revealed that while certain learners are able to compartmentalize the cognitive processes involved in task performance, break it all up into different stages, and identify what they can express linguistically, for others this may be more challenging. The current dissertation supports previous literature (Gurzynski-Weiss & Révész, 2016;
Sasayama, 2016;) in the importance of assessing L2 language proficiency and examining its role in task complexity research.

6.2 Research Question 2a. Relationship between Cognitive Task Complexity and Levels of State Anxiety as Experienced during Task Performance on the Two Levels of Cognitive Complexity

The first main hypothesis that led RQ2 was empirically supported in the results. Specifically, it was predicted that as cognitive task complexity increases, state anxiety would also increase due to the impossibility to attend to multiple cognitive processes occurring simultaneously. This hypothesis has been specifically addressed in RQ4 in an exploration of the three-way interface Language Anxiety—Cognition—L2 Performance.

There is a paucity of studies to date that have investigated the relationship between cognitive task complexity and language state anxiety during task performance by using a measure that actually attempted to capture the temporary and fluctuating nature of anxiety. Only one study so far, Baralt and Gurzynski-Weiss (2011) addressed this problematic issue and developed a scale that examined the levels of anxiety concurrently to task performance and administered it at two points in time during the task. The study investigated whether state anxiety levels varied across type of medium—CMC versus FTF in two interactive tasks. Although the researchers found no statistical differences between computer mediated communication and face-to-face on levels of state anxiety, one of the unique contributions of the study was the creation of an instrument that examined the dynamic nature of anxiety per se. In this regard, the current study contributed to this strand of research within TBLT that took one step further in the conceptualization of the construct of language anxiety to look at its temporal dynamic nature in L2 tasks. The current study took the study of
language anxiety one step further and investigated Zeidner and Matthews’ three facets of language anxiety: affective/somatic, cognitive and behavioral (which will be discussed in further detail in the discussion of RQ4).

To address the interplay between Language Anxiety – Cognition - Task performance, it was decided to manipulate cognitive task complexity. By manipulating cognitive demands in L2 tasks, the relationship between cognitive load and state anxiety could be investigated. As discussed in the literature review section, it has been argued that learners’ factors are more likely to play a role in more cognitively complex tasks over simpler ones (Robinson, 2007). In this regard, TBLT researchers (Robinson, 2011; Torres 2013) have highlighted that L2 performance and development are predicted to change, and this variation will strongly depend on individual differences.

Turning to the quantitative data, the main hypothesis was supported as levels of state anxiety increased as cognitive demands of the task did. Overall, the cognitively complex task yielded a meaningful increase in the levels of state anxiety as perceived by learners. In other words, statistical significant differences were observed between the two levels of cognitive task complexity. Not surprisingly, for the majority of participants (57%), when they had to attend to multiple cognitive demands during oral task performance, they perceived their anxiety levels to have risen significantly. On the contrary, self-ratings and qualitative data also revealed a trend that, for some of the participants 21(%), increasing the cognitive demands of tasks—employing the +/- Few Elements operationalization—in fact led to a reduction in participants’ levels of state anxiety (rather than an increase).
The current findings agree with Robinson (2007) and Kim and Tracy-Ventura (2011) regarding the relationship between foreign language anxiety and cognitively task complexity on L2 performance and L2 development. In accordance with Robinson (2007), results in the present dissertation overall suggest that more complex tasks lead to higher anxiety levels during task performance.

Among the most important findings, it is interesting to note that Robinson (2007) also employed self-assessment via the “input, processing and output anxiety scale” (IPOAS) (MacIntyre and Gardner, 1994) to investigate state anxiety across three levels of task complexity. Results of the study indicated that levels of anxiety were significantly higher in the output stage of task performance. These results appear to be consistent with findings in the current dissertation in that state anxiety in this case also increased statistically after task completion (as compared to Time 1, half-way through the task). In Robinson’s 2007, levels of anxiety were reported by participants to be lower during the input and processing stage and significantly higher at the output stage. Importantly, whereas Robinson (2007) utilized interactive oral L2 tasks on Japanese learners of English, the current study used monologic narrative tasks.

These findings seem to indicate that the pervasive effect of state language anxiety during L2 task performance affects interactive as well as monologic tasks, and appears to be present in various languages, such as, L2 English (Robinson, 2001, 2007; Tracy-Ventura, 2011) and L2 Spanish (current study).

As mentioned in the literature review section, neither Robinson or Kim and Tracy-Ventura clearly defined how language anxiety was being operationalized. On the one hand, despite Robinson partially addressing the temporal dynamic nature of state anxiety by
using the IPOAS scale, its conceptualization of language anxiety seems restricted to the assessment of anxiety at the input, processing and output stage. In his conceptualization of language anxiety, the personality trait dimensionality, however, was not considered, and thus, the view of language anxiety as concurrent to task performance appears too simplistic.

Results in the current dissertation further supported Tracy-Ventura (2011) in learners’ perceptions of the negative effect of language anxiety on their task performance. Nonetheless, the author employed a trait measure to measure trait anxiety as concurrent to task performance, and thus, his operationalization of anxiety is problematic.

The current findings also agree with Gilabert et al. (2009) and Robinson (2001), who claimed that language anxiety had a negative impact on learners’ perceptions of task difficulty on more cognitive complex tasks versus the cognitively simpler tasks.

Lastly, contradictory results seem to exist, nevertheless, with Révész (2011)—who found no meaningful differences that language anxiety did not seem to affect task performance. The author attributed no differences between task complexity levels to the development of coping skills on the part of the learners.

**6.2.1 Major Factors Influencing Language State Anxiety during Oral Task Performance**

Among the most enlightening findings were the qualitative data that explained the main reasons for these two patterns. First, resource-directing resources (Robinson, 2001), specifically +/− Few elements were found to be a significant and determining factor affecting learners state anxiety. Interestingly, it was found to be related to debilitative and facilitative state anxiety.
As discussed in the Results chapter and in the discussion of RQ1, participants appeared to have experienced the presence of more action and higher number of elements to be distracting and stressful. So many characters that looked similar, the complexity of the pictures and the more complicated storyline made them feel overwhelmed and that was stressful and nerve-wrecking as they had to do too many things at once. It is not surprising that, for the majority of these learners, having to think about a coherent story, use vocabulary and grammar to use to describe the action, and produce it all with creativity in such little amount of time was extremely anxiety-provoking.

However, participants’ comments also highlighted the opposite trend. For these learners, data revealed that cognitive task complexity does not necessarily result in higher levels of state anxiety in the learner. Particularly for low proficiency learners, increasing conceptual information led to a decrease in their anxiety levels, as learners had more to talk about in the complex task. For this group of learners, being able to focus on a wider variety of elements (portrayed in the pictures) aided their linguistic performance in Spanish. These results also are in accordance with the findings on the current dissertation on task difficulty (RQ1). The reader is reminded that for the task difficulty self-ratings, whenever participants could not narrate what was happening to a particular character due to a lack of words, they could simply describe the action in relation to other characters or speak about background elements. This finding was consistent with the results on task difficulty from the current dissertation discussed above, and previous task complexity literature (Sasayama, 2016; Gurzynski-Weiss & Révész, 2016).

Aside from conceptual information or task-internal factors being revealed as an important source of language anxiety, other factors contributing to state anxiety were
shown in the qualitative data. Again, supporting Gurzynski-Weiss & Révész (2016), results indicated at lower levels of proficiency linguistic task demands seem to matter the most, as the source contributing in a greater degree to levels of language state anxiety.

For these particular low proficiency learners, the current study seems to be more consistent also with Skehan’s (1998) Limited Attentional Capacity Model (1998) and Ellis (2003) in his conceptualizations of task complexity, since both models attempt to explain the role that task linguistic-demands play in relation to the cognitive demands of L2 tasks. While both Skehan and Ellis consider this component in their models, Robinson (2001, 2011) does not.

Not surprisingly, task mode or the oral nature of the task was revealed as the third most mentioned problematic factor inducing anxiety. The results from the current study also go in accordance with the most SLA literature on language anxiety that found oral anxiety to be more often and more common in L2 learning over writing anxiety. Evidence exists today that although language anxiety has been shown to have a detrimental effect in all language skills (listening, speaking, reading, and writing anxiety), it is the oral production that has been most widely investigated and regarded as the most anxiety-inducing skill over the other linguistic skills (Chen et al., 1999; Horwitz et al. 1986, Gkonou, 2013; Mak, 2011; Woodrow, 2006).

Aside from the level of complexity, the type of task—a narrative versus an interactive task—was proved to have a major impact on the learners’ state anxiety. The type of task strongly determines the type of language elicited in the L2. In the current study, an oral monologic task (as opposed to an interactive task), was found to be an important factor leading to the learners’ language anxiety during performance.
In this regard, the format of a picture description is associated with a very specific topic, which implies the use of specific L2 vocabulary and grammar structures associated with the story learners have to narrate. The complexity of the story, and the conceptual information determine the complexity of the code (Skehan, 1998) that learners are required to use in the L2. Specifically, for these particular learners of a lower proficiency level, learners’ state anxiety seemed to have increased due to a lack of an interlocutor to provide feedback. In that sense, it was reported by a vast number of participants that a major challenge in learners’ oral production was 1) not having a partner in a more natural oral interaction and 2) the main linguistic demands associated with the content of the story were the major source of anxiety, and they were directly linked to the type of task.

Although the current dissertation focused on factors related to the task, importantly, the effects of both task-internal and task-external factors on language anxiety were revealed in participants’ comments, and therefore, seem to have operated simultaneously during task performance (Robinson, 2005).

Among some of the factors external to the task most commonly mentioned by participants were the presence of a native speaker, the presence of other peers in the room, time pressure to finish the task, and environmental factors (such as noises and the presence of other people surrounding them). When looking at common factors playing a role in previous foreign language literature in SLA, Gkonou (2013) reported fear of speaking in front of peers and the teacher and the classroom setting as the most severe threats.

The current findings were in accordance with previous literature (Rubio-Alcalá in Gkonou, Daubney & Dewaele, 2016) who found how low confidence and self-esteem in the target language (TL) are directly linked with language anxiety. In this regard,
occurrences of low-self-esteem and low confidence, very much related to the learners’ self-concept and identity in Spanish as their L2, was reflected—explicitly or implicitly—in the qualitative data by approximately 85% of all participants. Interestingly, many of the participants’ comments seemed to agree with Rubio-Alcalá’s (2016) findings of the strong link between low self-esteem and L2 confidence and language anxiety.

In this regard, especially important in relation to language state anxiety, the current study suggests that both task-internal factors must be carefully considered in combination with factors external to the task. Particularly, situational factors, such as, task conditions and interactive factors play a major role in connection to learners’ affective factors during task performance.

6.3 Research Question 2b. On the Relationship between State Anxiety and Foreign Language Classroom Anxiety

As part of RQ2, and in an attempt to offer a more inclusive conceptualization of language anxiety, the current dissertation explored the trait-state anxiety dimensionality within the learner under a TBLT perspective. Surprisingly, to date, no SLA study or TBLT research exists that has investigated these two types of anxieties together by employing a TBLT approach.

A task is an ideal tool to study this relationship between trait and state anxiety as it allows language instructors, curriculum developers and SLA researchers to manipulate factors of the task that can potentially explain how state anxiety varies while learners are performing a particular task, and importantly what factors—and to what extent these factors—impact levels of language anxiety during task performance. A task—viewed as communicate outcomes (Swales, 2009)—is the perfect scenario to explore the relationship
between learners’ proneness to anxiety (trait anxiety) and state anxiety in the classroom setting, as it helps us further evaluate how state anxiety is affected in an instructional setting and how pedagogical materials should be designed in order to aid learners reduce and adjust affective factors.

In the current study, trait anxiety was operationalized as classroom anxiety rather than using a psychology instrument, like the state-trait anxiety inventory (STAI) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) or any other existing scale used to explore the individual’s trait anxiety. Investigating foreign language classroom language anxiety in the present dissertation (as trait anxiety) allowed us to interpret state anxiety scores specifically in a foreign language instructional setting. In addition, since the standard measure to assess language anxiety in SLA has been Horwitz et al.’s (1986) FLCAS, findings in the current study could be compared to previous literature. Furthermore, classroom anxiety was used a baseline to interpret and understand learners’ anxiety in both task complexity levels.

It was speculated that if the learner’s personality showed a higher predisposition to feel anxiety in the foreign classroom, then potentially, his/her levels of state anxiety during task performance will be higher than for a learner who shows a lower predisposition for language anxiety in the Spanish classroom. Likewise, it was hypothesized that a learner who is more prone to language anxiety may show higher levels of state anxiety when performing a complex task (over a simpler one) than a learner who is biologically less prone to anxiety and also needs to perform a task that imposes higher cognitive demands.

Interestingly, the current study found a lack of relationship between the classroom anxiety and state anxiety scores both at Time 1 and Time 2. The weak correlations between
learners’ predispositions to feel anxious did not seem to be associated with state anxiety during their oral performance, at least for these particular learners and for these specific tasks. Importantly, results from the current dissertation appear to support Spielberg’s (1966) understanding of trait-state anxiety. In this theory, Spielberg claimed that although state anxiety was dependent on the individual’s proneness to anxiety, not all people that display high levels of trait anxiety will necessarily experience high levels of state anxiety (in Zeidner & Matthews, 2011). This may be explained because of the multiple factors that intervene in the language leaning process, which makes it again very difficulty to tease out the extent to which these anxieties affect each other.

In addition, this trend could also indicate that these constructs may have to be viewed as separate independent phenomena. Importantly, classroom anxiety was defined by Horwitz et al. (1986) according to three main components: ‘fear of negative evaluation’, ‘oral anxiety’, and ‘test anxiety’. However, by definition these components appear to be different phenomena and may not necessarily explain learners’ temporary responses when cognitive task complexity is manipulated. This seems logical if we consider that a task does not have the same characteristics of a test. Moreover, although fear of negative evaluation in the Spanish classroom may have been presented during task performance for some of the participants, it may not have been present for other participants. In this respect, a significant number of participants in the qualitative data reported that stakes while conducting the experiment were low as they knew that their performance would not have any consequence in their course grades. In this sense, the actual setting where the study took place may have explained this lack of relationship trait-state as scores measured by
the FLCAS for language classroom anxiety could have been different in an actual classroom setting.

6.4 Research Question 3. On the Dynamicity of Language Anxiety Over Time

Language state anxiety—as perceived by learners—was revealed to fluctuate significantly over time—from Time 1 to Time 2 on the two-presumed cognitive task complexity levels (RQ3). As hypothesized, participants seemed to have experienced statistically significant differences on the levels of language anxiety from one point of the task to the other. Contrary to expectation, state anxiety—as perceived by learners—increased after task completion when compared to it level captured right in the middle of the task. This contradicts the results from Baralt and Gurzynski-Weiss (2011), the only TBLT study to date that attempted to operationalize state anxiety and also measured it at different points in time during task performance. Among the most relevant findings, the researchers found a slight decrease from Time 1 to Time 2, but this difference was not statistically significant. In other words, the study concluded that levels of anxiety were higher half-way through the task and then decreased after task completion. The current findings, however, seemed to point to contradictory results, corroborating that levels of state anxiety as perceived by learners indeed fluctuate over time.

Importantly, in agreement with Gkonou et al. (2016), this study supports the recent dynamic turn in SLA and hopes to contribute to shed light to the “increasingly influential paradigm shift” (p. 2) taking place in SLA research on language anxiety. As Dewaele (2012) argued in relation to the study of affective factors, theoretically and empirically, the field of SLA has started to view language anxiety under a ‘process-oriented’ perspective (p. 43). Following this view, recent studies in SLA (e.g. Dewaele, 2012; Gkonou et al.,
(2016) and TBLT research (Baralt & Gurzynski-Weiss, 2011) have highlighted that affective factors must be conceptualized as dynamic variables that evolve over time (rather than static phenomena), and this variation is contingent upon a multitude of contextual variables in constant interconnectedness. It is essential to point out that this ‘dynamic turn’ in the study of affective variables has emerged to a great extent due to the influence of the Complex Dynamic Systems Theory (CDST) (de Bot et al, 2007; Larsen-Freeman & Cameron, 2008) discussed above. The current dissertation recommends and supports the idea that future investigations on language anxiety should follow a conceptualization of state anxiety that applies a Complex Systems Approach.

In this respect, the evidence shown in RQ4 on the interplay Language Anxiety – Cognition – L2 Performance further supports that state language anxiety should be conceptualized as both the cause and consequence of the same process simultaneously during the same L2 event. This pattern sheds light on the controversy posed in previous SLA literature on the ongoing debate on whether FLA should be viewed as the cause or the consequence of L2 achievement (Horwitz, 2001). The current findings highlight the obvious limitations of the great body of past research that questioned the directionality of the relationship between FLA and L2 achievement. In this regard, results provide further evidence that the cause-consequence debate is fundamentally limited in its theoretical conceptualization of language anxiety, as it is impossible to capture the nature of affective factors in a one-way relationship rather than a two-way dynamic relationship.
6.5 Research Question 4. On the Three - Way Interface Language Anxiety—Cognition—Second Language Performance: The Three Facets on Foreign Language Anxiety from a Task-Based Language Teaching Perspective

The second main prediction that motivated the current study and RQ4 was an investigation of the interplay Language Anxiety—Cognition—Task Performance in an instructed SLA context. As expected, the current study provided evidence that state anxiety—as concurrent to task performance—seems to have a major negative effect on a variety of cognitive processes (RQ4a), and that those, subsequently appear to be perceived by learners as detrimental on their task performance (RQ4b).

Qualitative data obtained from the participants’ responses seem to go in line with the main findings from RQ2, where increasing the cognitive load in the complex task for the majority of participants also led to a debilitative type of anxiety. One of the most important contributions of the present study has been to address the three facets of anxiety: the affective/somatic, the cognitive and the behavioral facet. The psychologists Zeidner and Matthews (2011) defined these facets as follows: The cognitive facet of anxiety is “related to how information is processed in stressful situations” (p. 15). With regard to this facet, “anxiety is usually accompanied by changes in both the cognitive processes themselves (e.g., heightened attention to threat, worry), and in the contents of cognition (the particular threats about which the person is thinking)” (p. 15). ‘Worry’ (Sarason, 1988) is considered in psychology as the most powerful cognitive component of anxiety. In the second place, the authors claimed that the affective or somatic facet was associated to all “objective feelings of physiological arousal, and the more subjective emotions, such as bodily tensions, activation of the sympathetic nervous system, making the body ready for
bursts of energy” (p. 15). Lastly, the behavioral facet refers to defensive behaviors such as escaping or avoiding threat (Zeidner & Matthews, 2011). As discussed by the authors, what is interesting is that, although in the case of animals, these behavioral responses are usually “freezing or fight back” (when cornered by another animal or danger), “for people, they are easier to override through cognitive control of behavior” (p. 15). This control of behavior is coping strategies human beings have developed to overcome anxiety as a biological response.

Results from participants’ self-ratings and qualitative data provided evidence for the interconnectedness of these three components, feeding into each other, demonstrating a constant and cyclical relationship between the three facets of anxiety. On the qualitative data, 72% of participants showed the effects of language anxiety on cognition and on task performance.

As part of this dynamic relationship, it was further revealed how disturbances in the thought process induced state anxiety—viewed as the consequence of a decline in the learner’s processing. Results further suggested that the resulting state anxiety, in turn, continued leading to slower processing, in this case as the causal element of the equation. As discussed in the results section, participants’ comments showed L2 performance has been shown to constantly affect and being affected by the other two components during task performance. Reflecting this trend, the results presented here seem to support the hypothesis proposed by the adapted model discussed in the literature section (Donate, 2018).

In relation to the most important cognitive processes affected, a clear trend emerged in the results regarding the cognitive processes highly mentioned in participants’
comments. It was found that the most highly impacted cognitive processes affected by language anxiety were: ‘attention’ (Eysenck, 1992; Derakshan & Eysenck, 2009), ‘concentration/focus’, and ‘general thought process and idea organization’. In connection to the cognitive facet of the learner, the current findings further supported Gkonou (2013) who found ‘fear of making grammatical errors’ and ‘difficulties in producing oral output’ were commonly mentioned as one of learners’ most frequently mentioned cognitive disturbances connected to language anxiety. Learners’ worry of poor L2 performance seems to agree with Zeidner and Matthews’ (2011) idea of ‘worry about a present threat’ as one of the main reflections of the cognitive facet of anxiety.

This finding further supports the little evidence existing today in SLA (Chen & Chang, 2009; MacIntyre & Gardner, 1989, MacIntyre & Gardner, 1994a, Steinberg & Horwitz, 1986) that attempted to illuminate on the direct relationship between language anxiety and cognitive demands. These studies have shown that language anxiety has a negative effect cognitive processing, and subsequently, L2 task performance. From a TBLT perspective, barely few researchers (Robinson (2007) have emphasized the need to examine the potential effect of FLA on learners’ cognitive demands, and subsequent L2 linguistic outcomes.

The current results are also consistent with cognitive psychology research (Eysenck, 1992; Derakshan & Eysenck, 2009; Eysenck, Derakshan, Santos & Calvo, 2007) that has addressed the effects of anxiety on cognitive processing and task performance. By applying attentional control theory, these studies have employed eye-tracking methodologies to examine the negative impact of anxiety on cognition and performance. These findings provide further evidence to support the presumed interconnectedness
between language anxiety, cognition and L2 performance, which offers validation for the language anxiety dynamic model proposed in this dissertation and adapted from King and Ng (2018). A current proposal (Goetze, dissertation in progress) investigating the same interplay between teacher’s emotions (with a focus on language anxiety) and cognition further supports these findings from the teachers’ perspective. The study provides evidence of the close interrelationship between teachers’ emotions and cognitive processes.

6.5.1 Debilitative and Facilitative Anxiety

Results of the current dissertation have revealed that language anxiety can be both facilitative and debilitative, depending on the positive or negative effects on the individual’s performance. Importantly, participants’ comments appear to strongly imply that language anxiety in an instructional L2 context seem to be presented in both forms during the simple and the complex task. This pattern is in accordance with the psychologist Scovel (1978), who argued that both anxieties usually work together. Findings also go in line partially with Chastain (1975) and with Kleinmann (1977), who found positive correlations between classroom language anxiety and L2 achievement. Nonetheless, the authors’ examination of anxiety was limited to the relationship between classroom anxiety and final course grades, and hence, whether the state anxiety (as concurrent to task performance) may have a positive effect on task performance was never explored.

When comparing previous task complexity literature that examined the relationship between language anxiety and task performance, findings are mixed. Interestingly, the current study partially supported Révész (2011) in finding that the use of coping skills that explained the lack of a relationship between cognitive task complexity and language anxiety. In her study, no statistical differences on learners’ linguistic performance were
found between the cognitively simple and the cognitively complex tasks with respect to language anxiety. Robinson (2007) found linguistic benefits in learners’ performance on the task designed to impose higher cognitive demands. His findings confirmed the Cognition Hypothesis, as learners’ oral production showed more interaction, linguistic complexity and uptake over the mid-complex and simple tasks. A negative correlation was found between language anxiety and performance. Language anxiety was accounted for as part of task difficulty construct. The author concluded the low anxiety group produced more complex language in the more complex task than the high anxiety group (Robinson, 2007).

These operationalizations and conceptualizations of language anxiety as emotional temporal states seem to present limitations. As discussed above, it seems problematic that the studies above did not appear to account for the complexity of language anxiety as a construct.

In view of the present findings, I argue that the contradictory results of these studies are partly due to the fact that the complicated and multi-dynamic and multi-faceted nature of language anxiety (as an affective factor) has not been fully captured either in TBLT research or in previous SLA literature on FLA. Due to the limitations both methodological and conceptual, drawing conclusions about how language anxiety is connected to learners’ performance has become more difficult and highly speculative.

6.6 Research Question 5. On the Relationship between Cognitive Task Complexity and Second Language Performance Measures

RQ5 sought to provide an answer to how cognitive task complexity was related to L2 performance by employing CAF measures. To better address how language anxiety was
connected—and had either a debilitative or facilitative effect, or both—the current dissertation employed linguistic measures to find out if learners benefited as cognitive task complexity increased.

It is also part of the on-going debate on the effects of increasing the cognitive demands of tasks on L2 performance of the triad CAF. In accordance to Skehan’ LAC model, findings on the current study seemed to suggest a trade-off existed between learners’ syntactic complexity and accuracy when cognitive task demands where manipulated. Results showed that while learners were statistically and significantly more fluent and produced statistically and significantly more complex language on the complex over the simple task, they were, nevertheless, less accurate. In other words, overall participants were statistically more accurate on the simple task over the cognitively complex task. Contradicting Robinson’s (2001, 2005, 2011) CH, increasing task complexity did not seem to lead to a better performance on both learners’ accuracy and syntactic complexity.

Among the most important findings of the results, it is noteworthy the use of coping strategies in the complex task. The fact that participants’ linguistic fluency and syntactic complexity in the complex task improved—when compared to the simple task—seems to suggest that learners were able to overcome the language state anxiety while performing, and still produce a better story that was linguistically more complex and included more information.

The present study partially supports and partially disagrees with previous studies using CAF measures (Norris & Ortega, 2009; Sasayama, 2016). On the one hand, it is interesting to note that mixed patterns exist in relations with CAF measures and previous
literature. While Sasayama (2016) found low proficiency learners only benefitted in the syntactic complexity measures, in her study, no advantage was found for fluency and amount of language produced (only high proficiency levels benefitted regarding number of words produced). This dissertation, nevertheless, found a different pattern in that low proficiency learners did improve their task performance in all fluency and productivity measures.

Findings partially agree with Norris and Ortega (2009), who argued that amount of subordination tends to be apparent at lower levels of proficiency, particularly at lower levels of proficiency. It is interesting to note that for the current study, it was found that low proficient learners did seem to produce higher levels of global syntactic complexity on the cognitively complex task when compared to the cognitively simple task. Critically, results showed learners equally benefitted linguistically in their phrasal complexity, which commonly tends to improve at higher proficiency levels (Norris & Ortega, 2009).

For accuracy, whereas Sasayama’s findings revealed learners’ performance appeared to show no positive advantage on the simplest task (the alarm clock story), the current study revealed a clear advantage in the number and severity of the errors produced, showing a more accurate performance in the cognitively simple task. Future research should incorporate measures of lexical variety to examine if any advantages exist in different levels of task complexity as a result of language state anxiety.
CHAPTER 7: LIMITATIONS, FUTURE DIRECTIONS AND PEDAGOGICAL IMPLICATIONS

The results presented in the current study have made valuable contributions to the ISLA and TBLT strands of research in relation to the role of language anxiety in L2 pedagogic tasks. Nonetheless, it is important to consider the main limitations of this investigation as well as the areas that deserve further future examination in order to accurately interpret the patterns observed. In the following section, the main limitations, further directions and recommendations for teaching practice are discussed.

7.1 Limitations

The current dissertation attempted to shed light on the multifaceted and dynamic nature of language anxiety in an L2 context utilizing a TBLT approach. Although it is hoped that a contribution has been made to better understand these dimensions within the L2 learner, it is important to point out several limitations (besides those already mentioned in the discussion chapter) presented in the current study.

The first one is that the present dissertation mainly employed subjective indirect measures of cognitive load, namely, task difficulty and mental effort. However, as pointed out by Brünken et al. (2003), it is necessary to take into account that these measures (as most self-report measures do) pose certain limitations in that a direct causal relationship cannot really be assessed between the results of the measure and the cognitive load imposed by them. Although the decision to use these measures was based on the large effect sizes found by Sasayama (2016) on task self-assessed task difficulty and self-perceived mental effort, employing a wider variety of cognitive load measures (in connection with state anxiety measures) would provide a better insight of the relationship between language anxiety as concurrent to task performance and cognitive processing. Importantly, results
from the present study further suggest that cognitive load measures must be carefully
selected according to the type of task and the level of proficiency. Following recent
recommendations (Sasayama, 2016), triangulation of subjective and objective measures is
highly recommended to examine cognitive load as a construct, and to validate cognitive
task complexity at low levels of proficiency. In addition, the inconsistency observed in the
findings of the previous task complexity research regarding task difficulty is not surprising
considering that the vast majority of these studies have employed very different
operationalizations of the construct of task complexity. Due to this inconsistency and its
multiple operationalizations existing today in the literature, it has become a major
challenge for TBLT researchers to establish comparisons accurately and make results
interpretable.

As for cognitive load, the language anxiety measures employed in the present
study—subjective indirect measures—pose similar challenges in investigating the role of
language anxiety overall. On one hand, it is important to note that the FLCAS poses serious
methodological limitations in its approach to account for language anxiety. Despite being
considered the standard measure of language classroom anxiety until today, in the past
decade, SLA researchers have started to point out that its ‘quantitative methodologically
and largely quantitative approach’ (Gkonou et al., 2016, p. 4) has become problematic. The
main issues attributed to this measure are related to its difficulty to account for the
multitude of variables influencing the behavior and variability of language anxiety.

Another limitation of the current study was that it was conducted in a laboratory
setting, which may have affected learners’ levels of anxiety differently than in an actual
instructional setting. Therefore, TBLT and SLA researchers in the future should consider
conducting studies in the actual classroom setting, which is associated with curricular issues, such as grades (see Leow, in press for a curricular approach to ISLA research). Whereas, the lab setting was actually suggested by some participants to be a predictor of language anxiety, for others, it was not a source of anxiety. For these participants, stakes are higher in the classroom because a grade is given, whereas in the laboratory, participants had been informed their grade would not be affected.

7.2. Future Directions

In light of the current results and considering recent findings, it is strongly recommended to conduct further research on task complexity controlling for task type and for the particular operationalization employed by researcher. In this regard, the cognitive processes involved when manipulating resource-directing (e.g. +/- Few Elements,) versus depleting (e.g. +/- provision of pre-task planning time) (Robinson, 2001, 2005, 2011) appear to involve distinct cognitive processes when learners perform a picture description monologic task as compared to an argumentative or a reasoning interactive task.

In connection to the language anxiety measures utilized, future studies investigating state anxiety should administer the state anxiety measures more than twice during task performance when exploring its dynamicity to achieve a more accurate picture of its fluctuation over time. It is recommended that SLA researchers triangulate research designs employing a wider variety of both direct and more objective measures in combination with indirect measures to shed more light on the role of state anxiety during task performance.

For example, future research investigating language anxiety could further benefit from the use of other introspective methodologies, such as concurrent think-aloud (Leow, 2015) protocols to examine learners’ processes, or stimulated recalls to get a better insight
of the relationship between anxiety and processing. On the other hand, besides using subjective measures, researchers should triangulate research designs employing physiological measures used today in psychology to examine anxiety, such as skin response or electrodermal activity measures (Reimer, Mehler, Coughlin, Godfrey, & Tan, 2009; Shi, Ruiz, Taib, Choi, & Chen, 2007) heart rate (Paas & van Merriënboer, 1994), and eye-tracking measures (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Palinko, Kun, Shyrokov, & Heeman, 2010) to inform ISLA research further on the relationship between anxiety and cognition in the L2 and FL context.

In the future, it would be interesting to conduct more classroom-based studies for better ecological validity and to examine this relationship in a classroom context. As argued above, an important contribution of the current dissertation for the SLA and TBLT strands of research, has been the importance of distinguishing classroom anxiety (at least, as conceptualized by Horwitz et al., 1986) and state anxiety as separate constructs. Nevertheless, future SLA researchers investigating state anxiety should consider employing other measures from cognitive psychology to investigate the trait-state relationship, such as the well-known instrument used in psychology that specifically addresses the trait-state dimensionality, the STAI scale (Spielberg et al, 1983).

In the future, more research needs to be conducted examining other learners’ individual factors, such as working memory, language aptitude, motivation, personality type, language attitude, etc., in connection to language anxiety from a TBLT perspective. This would shed light on what additional individual differences affecting learners’ state anxiety during task performance. In this regard, task-internal and task-external factors should be analyzed together as they are in constant interaction. In this respect, it would be
worth investigating the effects of limited proficiency and cognitive task complexity, and whether a potential interaction exists on cognitive load and language state anxiety during task performance.

Based on Spielberg’s model of state anxiety, future SLA researchers should examine the relationship between trait anxiety and state anxiety by manipulating the stressor (stressful situation) to which the language learners are exposed. Importantly, the facilitative-debilitating roles of language anxiety, and the three facets of anxiety (affective, cognitive and behavioral) (Zeidner & Matthews, 2011) should be explored in relation to: a) other types of tasks (monologic vs interactive), across task modes (oral versus written production), and in different learners’ populations (L2 learners vs heritage learners) for a better understanding of the nature of language state anxiety in a variety of L2 contexts.

In addition to other individual differences, future studies need to investigate the influence of other affective factors, besides language anxiety, to achieve a fuller understanding of the role of affect in language performance, development and acquisition. Following late proposals in the field from positive psychologists (Fredickson, 1992; Oxford, 2016; Rubio-Alcala, 2016) it becomes crucial to ISLA researchers to conduct empirical studies examining more rigorously the role positive emotions in connection with the negative ones. For over four decades, the ISLA and TBLT areas of research have focused on the detrimental effects that language anxiety and other negative affective factors have on L2 production and development. Nevertheless, recent psychologists (Fredickson, 1992; Oxford, 1990, 2016) have started to explore positive emotions in L2 contexts as a necessary step to explain their compensatory role for the negative emotions in language.
learners. It is argued in the current dissertation that it is essential to continue investigating positive affect to move the ISLA field forward.

In addition, future research should investigate language anxiety and other affective factors from a social perspective, taking into account factors external to the learner. As discussed in the literature review section, recent theoretical proposals such as King and Ng (2018) (in Mercer and Kostoulas, 2018) have evaluated affective factors from a dynamic perspective that incorporate a variety of social and contextual factors into the equation. Among these factors are the interaction between the teacher and the student, the learners’ cultural and institutional values, and other relevant social and contextual factors that are predicted to have a direct and major impact on learners’ L2 achievement. Affective factors are often determined by the individual’s cultures and thus, exploring the cultural, societal and institutional components is deemed essential to achieve a better understanding of the role of language anxiety.

Lastly, in relation to the linguistic measures, future studies should examine the effect of language state anxiety—as cognitive task demands are manipulated—on other linguistic measures, such as lexis. Participants made reference to how the anxiety experienced during task performance made them forget vocabulary and lexical items they usually knew. Therefore, measures of lexical diversity and lexical variety should be considered to find out whether and how the use of lexis is influenced by affective factors.

An important contribution of the current study has been to provide new evidence to the task complexity and the TBLT body of literature on the application of this newly introduced measure of accuracy at lower levels of language proficiency. As argued by Foster and Wigglesworth (2016), this accuracy measure has the advantage of evaluating
error gravity, something the majority of global and local accuracy measures tend to ignore. Nonetheless, it is recommended that future studies employ Foster and Wigglesworth’s (2016) weighted clause ratio (WCR) to test the efficiency of this measure across different proficiency levels.

7.3 Implications for Teaching and the Application Positive Psychology: Why Does It All Matter?

The current study presents several implications for teaching to be considered by language instructors, and for syllabus and curriculum development. Although the current study has focused mainly on task design features, findings have clearly indicated the relevance of a multitude of factors—both task related and task external factor—that seem to predict language anxiety (to a smaller or greater degree) in L2 contexts.

The ultimate goal of investigating language anxiety in the current dissertation is with the purpose of aiding language instructors and curriculum developers: gain awareness of what task factors and what learners’ factors play a major role in helping learners reduce their anxiety levels during task performance in an instructional context; to better understand the interconnectedness between emotions, cognition and how there are in constant interrelation; and to make better decisions in relation to pedagogic task design. Gaining a better understanding on how emotions play a role in the L2, and in what areas seem to have a major effect is of paramount importance to be able to design pedagogic materials and tasks that address learners’ emotional needs to promote L2 performance, development and, eventually, L2 learning.
7.3.1 Pedagogical Considerations on Task Design for Low Proficiency Levels

Among the most important implications to be considered for the language classroom evidenced by the current study is that designing tasks that are cognitive more complex may not necessarily translate into higher difficulty. In this respect, it is especially important that language instructors take into account the influence of the role of language proficiency when manipulating the cognitive demands of pedagogic materials. This observation seems to be particularly relevant for narrative picture-description tasks, results suggest that increasing the number of elements in a story will sometimes result into a more structured storyline and providing more opportunities for learners to build a coherent story. On the contrary, especially for low proficiency speakers, a simpler task with fewer elements depicted may be more challenging, as a less-structured storyline with less conceptual information may offer fewer chances for communication. In this case, learners may need to concentrate more and probably exert greater mental effort in comparison to a task that contains more elements or steps involved. For language teaching, this implies that a more complex task manipulated via resource-directing (Robinson, 2001) (at least for the +/- Few elements operationalization) seems to enhance learners’ performance by providing more structure to narrate a story in a creative way. The following excerpt reveals this idea:

*I didn't feel much different than from when I completed the prior exercise. The only difference is that there was more information provided in the pictures this time around which was both helpful and not because it was more restrictive but provided greater structure that I didn't have to think about and plan myself* (participant 28, complex task).

In addition, current findings appear to indicate that language instructors teaching at lower levels of proficiency should focus on linguistic demands (Gurzynski-Weiss & Révész,
2016)—which seems to be more important for novice learners—over higher proficiency levels, where cognitive processes seem to play a fundamental role in relation to task difficulty.

7.3.2 Positive Psychology to Inform Second Language Classroom Pedagogic Decisions: Raising Awareness on How Language Anxiety Actually Operates in the Second and Foreign Language

As discussed by Oxford (2016), although anticipatory anxiety usually has some advantages, for positive psychologists (Fredickson, 2001, 2003, 2004; Seligman, 2011), negative emotions “cause a narrowing of the individual’s response options to survival behaviors” (p. 181). In the past decade, positive psychologists have addressed positive emotions, such as, flow, agency, hope, optimism and positive emotions, while considered anxiety, anger and shame as negative emotions (Oxford, 2016). In a recent volume New Insights into Language Anxiety: Theory, Research and Educational Implications (edited by Gkonou, Daubney & Dewaele, 2016), Oxford reflected on how for pedagogical practice, it is strongly advised to rely on positive emotions in the foreign and second language context to compensate for negative emotions, such as language anxiety.

First, learners who inherently possess a higher predisposition for language anxiety in the foreign language classroom will tend to show very particular behavioral tendencies, symptoms and conduct that are very different from low-anxious learners. In this case, identifying which learners demonstrate higher levels of trait anxiety, and in what forms these are usually displayed, will potentially aid language instructors to look for more individualized solutions that help reduce language anxiety, and thus, enhance L2 development. Second, becoming aware of what particular stimulus—L2 tasks and
situations—tend to trigger learner anxiety most frequently becomes important for instructors to create the most adequate learning environment.

Since low self-esteem in the L2 is associated to trait (rather than state) anxiety, it is strongly recommended that language instructors make a conscious effort to identify those learners which show low self-esteem and low confidence in the L2 (Rubio-Alcalá, 2016) to be able to understand their needs better. Becoming aware of highly anxious learners and what their needs are will help us look for more individualized solutions and coping strategies to reduce their classroom anxiety and enhance L2 development and learning.

Learning about the three-way interface Language Anxiety—Cognition—L2 Task Performance, directly linked to the three facets of anxiety (Zeidner & Matthews, 2011): affective/somatic, cognitive and behavioral, will help language instructors understand how anxiety operates in a L2 instructional context, preparing them with better tools to recognize when learners show anxiety symptoms associated with anxiety. For example, for some learners, language anxiety will be demonstrated in the form of physiological reactions that may be visible to the teacher or peers surrounding them (e.g., sweating, difficulty breathing, certain facial gestures). However, for other L2 learners, language anxiety will be demonstrated as inability to focus, attentional disturbances or worry, or task avoidances. In these cases, visual signs are usually not shown and therefore, it will be much harder for instructors to note when the learner is experiencing anxiety. In the former case, anxiety is shown in its cognitive facet, and in the latter, as a behavioral tendency.

Importantly, language instructors can mistakenly think a learner is purposely not making a good effort in completing a particular task in the classroom, when it may be well that he or she simply is avoiding the task (typical tendency related to behavioral anxiety)
because he or she is feeling too anxious to perform it and hence, unconsciously avoids it. This is very typical of oral production tasks. In the instructional context, language instructors must be aware that certain L2 tasks and L2 situations—unique and specific to the language learning process—are more likely than others to induce higher levels of state anxiety. Some of these situations have been shown to be higher predictors of language anxiety, and thus, raising awareness will benefit learners and instructors equally.

The present study appears to provide evidence for how the complex network of factors that intervene in the second language learning process (Daubney, Dewaele & Gkonou, 2016), makes it extremely difficult especially for affective factors to often pinpoint the main causes for language anxiety.

Despite this challenge, more general tendencies in relation to what L2 situations and tasks will likely trigger anxiety also exist and can be applied to a majority of the language learners—irrespective of their level of proficiency and other individual differences. For instance, as evidenced in the current study and also in a vast number of previous SLA literature (Horwitz, 1986, 2001) oral and communicative anxiety (versus writing, reading or listening anxiety) are far more common and have been more frequently observed in instructional settings.

I would argue that another general tendency, reflected in the current study, is the potential positive role allowing time to plan prior to oral task performance may have on reducing learners’ state anxiety. Language instructors are advised to provide L2 learners, especially prior to tasks that involve oral production, with pre-task planning time to prevent language anxiety to arise. Especially important for low proficiency learners, a pre-task planning phase where special emphasis is made on practicing key lexicon (Newton,
—may not only affect learners’ perception of task difficulty (Gurzynski-Weiss & Révész, 2016), but more specifically to language anxiety. This may be explained on the premise that incorporating vocabulary building pre-task activities has been argued to ease the process of coding and decoding the linguistic message (Newton, 2001). Consistent with this idea is an unpublished study (Donate, 2015) conducted on the effects of planning time on learners’ emotions and perceptions on written production. Among the most relevant findings were that for the majority of learners planning time was perceived as a fundamental strategy to reduce their anxiety levels during task performance.

Lastly, it is helpful for language teachers to know that learner-centered methodologies are crucial, as is the use of interactive tasks and collaborative group work where the focus is on the learner, rather than the teacher. In this regard, it has been suggested (Rubio-Alcalá, 2016) that group or pair work has a positive effect on language anxiety as opposed to individual work or teacher-centered methodologies. As Rubio-Alcalá (2016) recommended, learner-centered methodologies usually help reduce language anxiety since “it gives the learners the opportunity be the main agents of the learning process” (Gkonou et al., 2016, p. 204). The author also pointed out that methodologies that are too centered on the instructor will more likely provoke feelings of anxiety in the learners, as usually they feel intimidated by the teacher, leading to fear of negative evaluation (Horwitz et al., 1986). The concept of agency is strongly related to the idea of self-concept and self-esteem, where learners are in charge of the own learning process and tasks allow them to explore and perform tasks on their own (Rubio-Alcalá, 2016). In connection to this, the current study supports the importance of the role of the teacher to mitigate anxiety in the L2 context. In the qualitative data, some participants mentioned
their supportive relationship with their instructor as a positive factor with often helped them feel better about their performance.

It is also recommended that language instructors establish a methodology based on rapport (Rubio-Alcalá, 2016) where they show care for their students and a positive bond is formed from the beginning. The results reported here provide further evidence for the importance to consider the affective role when designing pedagogic materials and tasks.

The current study highlights the importance to consider external factors to the task as apparent to some participants’ comments reporting on contextual factors, namely, the presence of a native speaker and other L2 learners in the room as anxiety-provoking elements. Turning to the classroom setting, it seems essential for language instructors to pay special attention to task design features (level of task complexity) and conditions in combination to situational factors during task performance to enhance the learning process. With regard to the latter, the current study strongly emphasizes that creating a positive classroom atmosphere based on rapport is deemed crucial to promote successful language performance. In addition, creating a positive environment should be applied in combination to developing effective classroom materials. Ensuring our role as language instructors (Price, 1991) promotes a relaxed and collaborative atmosphere appears to be of paramount importance particularly for those learners that show low self-esteem (Rubio, 2001, 2016) in the FL, because it strengthens their L2 self-concept. Language instructors are advised to apply these strategies that can help alleviate and relieve language anxiety in the L2 and FL classroom.
7.3.3 Implications for Assessing Learners’ Complexity, Accuracy and Fluency

Turning to the application of CAF measures by language instructors for teaching purposes, I would argue that when assessing syntactic complexity, the current findings seemed to emphasize the importance of employing a variety of measures that are able to evaluate learners on different aspects of syntax in the L2. Following Norris and Ortega (2009), examining global syntactic complexity only does not seem to be enough to offer language teachers a complete picture on learners’ abilities to use a wide range of sophisticated structures. Besides global complexity, utilizing a measure of subordination and clausal complexity appears to be more informative in informing language teachers of syntactic aspects that are important and constitute complementary areas of the learner’s syntactic linguistic competence.

In relation to accuracy, it is further advised that language instructors pay closer attention the severity of errors learners make, rather than only focusing on the number of errors made. When assessing language learners’ performance, it seems crucial to take into consideration whether the error made impedes communication or only constitutes a minor error that does not interfere with the meaning of the message conveyed.

Lastly, I argue that the use of both productivity and fluency measures are important to provide a more appropriate assessment of L2 performance in relation to learners’ L2 fluency. As for the syntactic complexity dimension, the amount of language produced (productivity) and speech rate (as gauged by syllables per second, and words per minute) also inform us of independent but complementary areas of L2 learners’ linguistic oral production. In this respect, future research employing CAF should examine both areas to provide a better picture of the learners’ language abilities.
CHAPTER 8: CONCLUSIONS

The current study sought to investigate 1) the relationship between cognitive load—examined by subjective and indirect independent measures—and cognitive task complexity with the purpose of validating presumed versus designed task complexity in L2 Spanish, 2) the relationship between cognitive task complexity and perceived state anxiety in oral task performance, 3) the relationship between perceived state anxiety and foreign classroom language anxiety (employed as baseline trait anxiety scores to interpret state anxiety scores in an instructional setting), 4) the dynamicity of language state anxiety—as perceived by learners—during oral task performance, 5) the interconnectedness of the three-way interface Language Anxiety—Cognition—L2 Task Performance, and 6) the relationship between cognitive task complexity and linguistic outcomes in L2 Spanish.

The first research question revealed that cognitive load was related to presumed cognitive task complexity for only one of the self-assessment independent measures, the mental effort measure. An important finding was that for lower levels of proficiency, perceived mental effort appeared to be a reliable measure showing large effect sizes in assessing cognitive complexity. For the task difficulty and the time estimation measures, no meaningful differences were found between the two levels of task complexity. One possible explanation for the lack of differences between the two tasks on task difficulty may have been due to the range employed on the scale (1-6 scale) of the measure, which may have not allowed for much variation in the participants self-ratings. Additionally, despite the past evidence supporting the reliability of task difficulty as a measure to address cognitive task complexity, the contradictory findings provide new evidence suggesting that
especially for low levels of proficiency, increasing cognitive demands by increasing the number of elements does not always translate into task difficulty. Supporting Sasayama (2016), the current study also showed time estimation does not seem to work for low proficiency learners of Spanish, as the measure did not seem to distinguish between the high and low distinct levels of task complexity. This underscores the importance that researchers need to consider the type of task and the level of proficiency very carefully when they design pedagogic tasks.

It may be argued that although only the mental effort measure (one out of three independent measures) showed statistical differences, the presumed cognitively complex task still appeared to have posed statistically significant higher task demands as compared to the simple task. Importantly, the difference in the number of characters between the simple and the complex tasks was 1 versus 9. The study supported Sasayama (2016) in that when employing +/- few elements, maintaining a large difference in the number of elements is necessary to keep cognitive demands statistically distinct. Another finding was that at lower levels of proficiency, task difficulty showed to be an independent construct from cognitive task complexity. With the exception of mental effort, additional cognitive load measures should be considered for L2 low proficiency speakers in future studies for validation of the construct, rather than only indirect subjective self-assessed measures of cognitive load. These seem to work better at higher levels of proficiency. Qualitative data revealed interesting findings on the relationship between presumed task complexity and the design features that seem to contribute to a greater degree to the actual cognitive demands. The linguistics demands, followed by the task’s conceptual information, were the task features that had a major role in the actual cognitive complexity.
The second research question showed that increasing the cognitive demands of tasks, via +/- few elements as the main task complexity operationalization, was related to higher levels of state anxiety, as perceived by learners participating in oral Spanish narratives. Importantly, a lack of relationship was revealed between learners’ state anxiety scores and their foreign language classroom anxiety scores, suggesting that these constructs are distinct and independent from each other. Future research is advised to assess this relationship in a wider variety of L2 tasks and L2 contexts to assess whether a relationship exists or not. SLA researchers should also employ a different trait measure to further test the *trait-state* dimensionality and find out if there is a relationship between the learner’s temporal emotional states and their predispositions to feel language anxiety in the classroom (trait anxiety). An important finding came out from participants’ comments on the factors affecting their language state anxiety to a greater degree. Among the task factors, linguistic demands—especially a lack of the right lexicon—task mode (oral anxiety), pre-task planning time, and the type of task were among those identified. Among learners’ internal factors, L2 self-confidence and low perceived proficiency were among the most important predictors that led to state anxiety during oral task performance.

The third research question showed that learners experienced statistically different levels of state anxiety when assessing their state anxiety over time. After task completion, learners’ state anxiety was perceived significantly higher (statistically) as compared to levels of state anxiety half-way through the task. In this regard, findings appear to suggest that the temporality and dynamicity of language state anxiety (as an affective factor in the L2) must indeed be carefully tested and tracked at different points in time during task performance to see how it fluctuates. In the future, this dynamicity should be examined as
an additional affective factor in the L2, in different types of task, task modes, and learners’ populations, particularly of different proficiency levels.

Turning to the interconnectedness between language state anxiety, cognitive processes and L2 linguistic task performance, crucial observations were found. Answering the first part of the fourth research question (RQ4a), participants’ qualitative data appears to indicate a strong link between state anxiety (as concurrent to task performance) and cognitive processes. The cognitive functions most generally reported by language anxiety were short-term/working memory, general thought process, thought process and idea organization, and focus/concentration.

The second part of the fourth research question (RQ4b), which addressed one of the main purposes of the current study, revealed a two-way cyclical relationship between language anxiety, learners’ cognitive processing and linguistic performance in pedagogic L2 tasks. Irrespective of the level of cognitive task complexity, this dynamic relationship seems to be constantly operating and maintained throughout L2 learners’ task performance. This finding was of paramount importance in supporting and validating the interconnectedness between emotional states, cognition and task performance in second language contexts.

Another insightful finding was related to the use of language anxiety coping skills by these low proficiency learners. In order to compensate for the detrimental effects of state anxiety, coping skills seem to be widely employed by learners to help them deal with the negative effects of the debilitating type of anxiety. The most reported coping skills were apparent by learners’ use of 1) task avoidance strategies and 2) emotional compensatory strategies to cope with low L2 self-esteem and L2 confidence (Rubio-Alcalá, 2016).
On the relationship between cognitive task complexity and L2 task performance, the present study showed that when learners performed the cognitively complex task, their performance overall improved in their oral fluency and syntactic complexity. On the other hand, for accuracy, learners showed a disadvantage on the cognitively complex oral narrative. Supporting Skehan’s (1998) model of limited attentional capacity, findings revealed a trade-off between syntactic complexity and accuracy.

In conclusion, the current dissertation hopes to contribute to the SLA field (more specifically the ISLA strand of research) and TBLT strand of research by providing insightful evidence on the role of language anxiety in L2 oral tasks. The present research is arguably the first study to investigate in-depth and in combination the core issues that characterize and define the nature of language anxiety from a TBLT perspective: the various task factors that seem to constitute major predictors of language anxiety during oral task performance, its relationship with cognitive task complexity, its dynamicity and temporal nature in L2 contexts, the different facets and dimensions of state anxiety, and its facilitative and debilitative roles in learners’ task performance in Spanish as a second language. This strand of research is a step forward in raising awareness on the importance that affective factors play in relation to L2 task design, material development, and L2 performance. It is also a step forward in emphasizing the importance of conducting interdisciplinary investigation, by contributing to inform the fields of TBLT, instructed SLA (ISLA), psychology, and second language education. It is hoped that future research continues exploring more rigorously the relationship between affective factors and cognitive processes in connection with L2 development and learning.
APPENDIX: MATERIALS

Appendix A: Language Background Questionnaire
(paper-based version of the original electronic language background measure administered via google forms)

Section 1

Email address ____________________________________

Section 2: Language Background Questionnaire

Biodata

Gender*

Male _____
Female _____
Other _____

Age*

________________________

Race/Ethnicity

White/Caucasian_____ 
Black or African American_____ 
Hispanic/Latino _____
Asian_____ 
American Indian or Alaska Native_____ 
Native Hawaiian or other Pacific Islander_____
Other____

**Section 3: Language Background Questionnaire**

What is your native language?*

English____

Spanish____

Other____

Country of origin

I was born in the US_____  
I was not born in the US_____  

**Section 4: Non-Native Language Background**

Please specify your country of origin

______________________________

Please specify at what age you moved to the US.

______________________________

1. Were you born in a Spanish-speaking country?

   Yes____    No____

2. Were your parents or grandparents born in a Spanish speaking country?
Yes_____  No_____ 

3. What language(s) do your parents/caregivers speak at home?

Section 5: Foreign Language Background

a) I am currently a…

  First-year student___  
  Second-year student___  
  Third-year student___  
  Fourth year-year student___  
  Graduate student___

b) In which Spanish class are you enrolled this semester at your University?

  I am not taking a class in Spanis right now.______  
  SPAN-021 Spanish Intermediate 1 _____  
  SPAN-022 Spanish Intermediate 2 _____  
  SPAN-103 Spanish Advanced 1 _____  
  SPAN-104 Spanish Advanced 2 _____  
  Other _____

c) Please specify the first language you have learned or studied.
Section 6: Spanish Language Background

a) How old were you when you first started learning/studying Spanish?

____________________________________

b) Please identify the context of instruction of Spanish (i.e. high school class, living abroad, etc.)

   Elementary school_____  Secondary school_____  Post-secondary school/University_____  Study abroad/living Abroad_____  Home language/language spoken by a parent_____  Other_____  

c) For how long have you studied Spanish? (i.e. 2 semesters in college, 3 weeks during summer camp)

____________________________________

d) Do you currently still receive formal instruction in Spanish?

   Yes____  No____
e) Please rate your proficiency in Spanish

<table>
<thead>
<tr>
<th></th>
<th>Novice</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Superior</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

f) Have you learned/studied an additional foreign language?

_____________________________________________________________________

Section 7: English Language Background

a) How old were you when you first started learning/studying English?

________________________________________

b) Please identify the context of instruction of English (i.e. high school class, living abroad, etc.)

Elementary school____

Secondary school____

Post-secondary school/University____
Study abroad/living Abroad____

Home language/language spoken by a parent____

Other____

c) For how long have you studied English? (i.e. 2 semesters in college, 3 weeks during summer camp)

_____________________________________________________________________

d) Do you currently still receive formal instruction in English?

Yes____        No____

e) Please rate your proficiency in English

<table>
<thead>
<tr>
<th>Speaking</th>
<th>Novice</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Superior</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F) Have you learned/studied an additional foreign language?

_____________________________________________________________________

290
Section 8: Other Foreign Language Background

a) How old were you when you first started learning/studying your first foreign language?

____________________________________

b) Please identify the context of instruction of your first foreign language (i.e. high school class, living abroad, etc.)

   Elementary school____
   Secondary school____
   Post-secondary school/University____
   Study abroad/living Abroad____
   Home language/language spoken by a parent____
   Other____

c) For how long have you studied your first foreign language? (i.e. 2 semesters in college, 3 weeks during summer camp)

   ______________________________________


d) Do you currently still receive formal instruction in your first foreign language?

   Yes____   No____

e) Please rate your proficiency in your first foreign language:

   Novice     Intermediate     Advanced     Superior     Native
<table>
<thead>
<tr>
<th></th>
<th>Speaking</th>
<th>Listening</th>
<th>Reading</th>
<th>Writing</th>
<th>Overall Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td></td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
<td>_______</td>
</tr>
<tr>
<td>F) Have you learned/studied an additional foreign language?</td>
<td>____________________________</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Please list below 1) what other languages you may have learned/studied, 2) how old were you when you started learning them? 3) what was the context of instruction? 4) how long did you study the language?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

Section 10: Experience Abroad in the Spanish-Speaking World

a) Have you ever been to a Spanish-speaking country/region?

Yes____  No____

b) If yes,

when_____________________________

where_____________________________

purpose of visit_____________________________

for how long_____________________________

c) Please include any information about your language use that you think may be important.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Appendix B: Task Instructions
(adapted from Sasayama, 2015)

Task Instructions

In pairs, you will engage in two story-telling tasks. For each story telling task, you will see a 6-cartoon strip on the handout. Based on the 6 pictures, please tell a story in Spanish.

STEP 1. First, you will have 30 seconds to look at the pictures and prepare your story. (NOTE-TAKING NOT ALLOWED)

STEP 2. After 30 seconds of preparation, please start telling the story as best as you can.

STEP 3. When you finish describing picture 3, please STOP TELLING YOUR STORY and raise your hand. The researcher will come and you will fill out a short questionnaire. Once you are done completing the questionnaire, continue telling the rest of the story where you left it (from picture 4 and until you are done.

STEP 4. When you finish telling your story, please stop talking and raise your hand to the researcher.

STEP 5. As soon as you finish the story, you will answer another questionnaire about the task you just performed.

STEP 6. When you finish the questionnaire, the second task will begin, and you will repeat the same process again.

Your story will be assessed based on its completeness (i.e., whether all six pictures are included in the story), effectiveness (i.e., whether the order of the
pictures is clear to the listener), and creativity; so, please do your best in telling the stories.

*Any questions?*

*I’d appreciate it if you could keep the content of this experiment confidential.*

*Thank you for your participation!*
Appendix C: State Anxiety Scale Mid-Task
(adapted from Baralt & Gurzynski-Weiss, 2011)

There is no right or wrong answer, so we ask you to simply provide your honest response in relation to this task.

Instructions: Please circle if you:

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

1. This task is stressful for me.
   6 5 4 3 2 1

2. I am not bothered by having to speak in Spanish so quickly.
   6 5 4 3 2 1

3. I feel more anxious in class than completing this task.
   6 5 4 3 2 1

4. I am getting flustered because I can’t process so much information and then talk about it.
   6 5 4 3 2 1

5. I don’t have enough time to figure out the story line before I have to talk.
   6 5 4 3 2 1

6. This task is fun and enjoyable.
   6 5 4 3 2 1
7. I feel tense having to talk in Spanish so quickly.

   6   5   4   3   2   1

8. This task makes me less anxious than I normally feel in class.

   6   5   4   3   2   1

9. I am relaxed and comfortable processing the information of this task and then talking about it.

   6   5   4   3   2   1

10. I feel like I have enough time to figure out the storyline before talking.

    6   5   4   3   2   1
Appendix D: Post-Task Anxiety Scale
(ratings’ section adapted from Baralt & Gurzynski-Weiss, 2011)

ID # ( )

Task # ( )

*There is no right or wrong answer, so we ask you to simply provide your honest response in relation to this task.*

Instructions: Please circle if you:

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

1. This task was stressful for me.

   6   5   4   3   2   1

2. I was not bothered by having to speak in Spanish so quickly.

   6   5   4   3   2   1
3. I felt more anxious in class than completing this task.

   6 5 4 3 2 1

4. I was getting flustered because I couldn’t process so much information and then talk about it.

   6 5 4 3 2 1

5. I didn’t have enough time to figure out the story line before I have to talk.

   6 5 4 3 2 1

6. This task was fun and enjoyable.

   6 5 4 3 2 1

7. I felt tense having to talk in Spanish so quickly.

   6 5 4 3 2 1

8. This task made me less anxious than I normally feel in class.

   6 5 4 3 2 1

9. I was relaxed and comfortable processing the information of this task and then talking about it.

   6 5 4 3 2 1
10. I felt like I have enough time to figure out the storyline before talking.

6 5 4 3 2 1

1.a. Please explain in detail all the feelings—good and bad—that this task provoked in you, and the reasons why you think you felt that way while performing.

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

1.b. How did that affect your performance in this task?

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

300
2. a. Please explain in detail what elements of the following task—storyline, characters, length, grammar you used, vocabulary, etc.—made the task more or less anxiety-provoking and how it did so.

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

2. b. How did that affect your performance in this task?

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________

______________________________________________________________________________________________________________________________
Appendix E: Self-Assessed Cognitive Load Measures  
(scale adapted from Sasayama, 2015)

**Task Difficulty**

3. a. Was the story telling task easy?

| VERY EASY | | VERY DIFFICULT |
|-----------|:|-----------------|
| 1         | 2   | 3               | 4               | 5               | 6               |

3. b. What made the story-telling task easy/difficult? Please provide details.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Mental Effort

<table>
<thead>
<tr>
<th>Very low</th>
<th>Very very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental effort</td>
<td>mental effort</td>
</tr>
</tbody>
</table>

4.a How much mental effort did you put in to complete this task?

1 2 3 4 5 6

4.b. What made the task require more or less mental effort? Please provide details.

_______________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________

_____________________________________________________________________________________________________________________
Appendix F: Cognitively Simple Task
(adopted from Sasayama, 2015; original source from Hill, 1960)
Appendix G: Cognitively Complex Task
(adopted from Sasayama, 2015; original source from Elder & Iwashita, 2005)
Appendix H: Retrospective Time Estimation Measure
(adopted from Sasayama, 2015)

Time Estimation

How long do you think your story was?

Indicate the approximate number of minutes and seconds you estimate to have spent in this task.

A) (        ) MINUTES

B) (        ) SECONDS.
Appendix I: Foreign Language Classroom Anxiety Scale (FLCAS) (Horwitz, Horwitz & Cope, 1986)

Instructions: Please circle if you Strongly Agree, Agree, Slightly Agree, Slightly Disagree, Disagree, or Strongly Disagree with each statement.

Note: There is no right or wrong answer. Your answers will only help researcher in second language to better understand learners’ feelings to the Spanish classroom, so we ask you to provide your honest response.

1. I never feel quite sure of myself when I am speaking in my Spanish language class.

2. I don't worry about making mistakes in the Spanish class.

3. I tremble when I know that I'm going to be called on in Spanish class.

4. It frightens me when I don't understand what the teacher is saying in Spanish.

5. It wouldn't bother me at all to take more foreign language classes other than Spanish.

6. During the Spanish class, I find myself thinking about things that have nothing to do with the course.
7. I keep thinking that the other students are better at Spanish than I am.

8. I am usually at ease during tests in my Spanish class.

9. I start to panic when I have to speak without preparation in the Spanish class.

10. I worry about the consequences of failing my Spanish language class.

11. I don't understand why some people get so upset over the Spanish class.

12. In the Spanish class, I can get so nervous I forget things I know.

13. It embarrasses me to volunteer answers in my Spanish class.

14. I would not be nervous speaking the Spanish language with native speakers.
15. I get upset when I don't understand what the teacher is correcting.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

16. Even if I am well prepared for the Spanish class, I feel anxious about it.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

17. I often feel like not going to my Spanish class.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

18. I feel confident when I speak in the Spanish class.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

19. I am afraid that my Spanish teacher is ready to correct every mistake I make.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

20. I can feel my heart pounding when I'm going to be called on in the Spanish class.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

21. The more I study for a Spanish test, the more confused I get.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

22. I don't feel pressure to prepare very well for the Spanish class.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
</tbody>
</table>

23. I always feel that the other students speak Spanish better than I do.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree Strongly</th>
</tr>
</thead>
</table>
24. I feel very self-conscious about speaking Spanish in front of other students.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

25. The Spanish class moves so quickly I worry about getting left behind.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

26. I feel more tense and nervous in Spanish than in my other classes.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

27. I get nervous and confused when I am speaking in my Spanish class.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

28. When I'm on my way to my Spanish class, I feel very sure and relaxed.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

29. I get nervous when I don't understand every word the Spanish teacher says.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

30. I feel overwhelmed by the number of rules you have to learn to speak Spanish.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

31. I am afraid that the other students will laugh at me when I speak Spanish.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

32. I would probably feel comfortable around native speakers of Spanish.
33. I get nervous when the Spanish teacher asks questions which I haven't prepared in advance.
REFERENCES


Sasayama, S., & Izumi, S. (2012). Effects of task complexity and pre-task planning on EFL


van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct


Yoshida, M. (2012). The interplay of processing task, text type, and proficiency in L2


