SECOND LANGUAGE LEARNERS' INDIVIDUAL DIFFERENCES AS A DYNAMIC SYSTEM: EVIDENCE FROM A RUSSIAN IMMERSION PROGRAM

By

Dmitrii Pastushenkov

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Second Language Studies – Doctor of Philosophy

ABSTRACT

SECOND LANGUAGE LEARNERS' INDIVIDUAL DIFFERENCES AS A DYNAMIC SYSTEM: EVIDENCE FROM A RUSSIAN IMMERSION PROGRAM

By

Dmitrii Pastushenkov

Guided by Dynamic Systems Theory (DST) and multidimensional perspectives on individual differences (IDs) in instructed second language (L2) learning (e.g., Larsen-Freeman, 2014; Sun & Zhang, 2020), this longitudinal mixed-methods study focuses on three conceptually different categories of L2 learners' IDs: working memory (WM), motivation, and the amount of L2 exposure. Despite previous propositions regarding cognitive, affective, and exposure-related IDs being interrelated (e.g., Pawlak, 2012), including WM and motivation (e.g., Serafini, 2017) and WM and the amount of L2 exposure (e.g., Denhovska et al., 2016), the longitudinal development of motivation and L2 exposure, as well as the effects of WM, motivation, and L2 exposure on learning gains in different skills and at different stages of L2 development have not yet been investigated. In this study, I aimed to address this gap and explore how L2 learners' motivation and the amount of L2 exposure change over time and how WM, motivation, and the amount of L2 exposure affect learning gains in lexicogrammar, speaking, and writing at different initial proficiency levels. Considering that Second Language Acquisition (SLA) and Instructed SLA (ISLA) research are dominated by studies of more commonly taught languages, I also strove to promote the need for more ecologically valid research with Less Commonly Taught Languages (LCTLs) and investigated IDs in the instructed L2 learning of Russian.

The participants were 52 students recruited from an eight-week Russian summer immersion program from four curricular levels (First-Year, Second-Year, Third-Year, and Fourth-Year Russian). The program was conducted online due to the COVID-19 pandemic. The results supported the view of L2 learners' motivation and L2 exposure as constructs showing stable and dynamic patterns and that different constituents of WM, motivation, L2 Russian exposure, and learning gains were engaged in complex relationships that varied at different curricular levels, language skills, and phases of the study. This dissertation includes a discussion of theoretical implications with regards to DST and the mixed-methods approach as future directions for ISLA studies of IDs, including research with LCTLs. The dissertation also includes a section on practical implications that discusses pedagogical aspects and implications for the development of background questionnaires and placement testing.

ACKNOWLEDGEMENTS

I would like to acknowledge my dissertation committee members, professors, and mentors Shawn Loewen, Aline Godfroid, Jason Merrill, Masatoshi Sato, and Paula Winke for their support over the course of my studies at Michigan State University. I am grateful for their assistance in the ideas that would become the crux of my dissertation. I would not be here without the help from my professors and mentors Phillip Hamrick, Ryan Miller, and Brian Baer from Kent State University and Irina Solovyova from Tver State University, who inspired me to explore the fields of Applied Linguistics and Second Language Acquisition and join the Ph.D. Program in Second Language Studies at Michigan State University. You believed in me from the very start of my studies, and I would never forget that.

I would also like to thank my wonderful colleagues and co-authors Cameron Camp, Caitlin Cornell, Evgeny Dengub, Irina Dubinina, Steven Garrison Gagnon, Curtis Green-Eneix, Olesya Kisselev, Tanya McIntyre, Rossina Soyan, and Iryna Zhuchenko, as well as my fellow STARTALK enthusiast Lizz Huntley. I am also very grateful for what I learned from my wonderful supervisors Alissa Cohen, Peter De Costa, India Plough, Charlene Polio, and Patti Spinner, to name a few. Working with you helped me become a better researcher and teacher.

Even though I cannot reveal the name of the summer program where I collected data for my dissertation, I am very grateful for all the help from the program's faculty and staff and their valuable insights into the program's philosophy and pedagogy. Visiting the program in person in the summer of 2019 was one of the most enriching experiences of my doctoral studies. I would like to thank the participants in my dissertation research for their time and effort and wish them

iv

luck in their future Russian language studies. This research was made possible through the *Language Learning* and NFMLTA / NCOLCTL dissertation grants.

Finally, I would like to thank my wife and colleague Olesia Pavlenko for her support over these years. I would never be able to reach this milestone with you. Thank you so much!

I would like to use this opportunity to encourage applied linguists to pursue graduate studies and open new horizons. The Second Language Studies Program at Michigan State University has opened a new world for me, and I will be always grateful for what I learned in this program.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	xi
KEY TO ABBREVIATIONS	xiii
CHAPTER 1: INTRODUCTION	1
Situating the study	1
Need for the study	3
Significance of the study	4
Organization of the dissertation chapters	6
CHAPTER 2: LITERATURE REVIEW	8
Dynamic systems and multidimensional perspectives on IDs	8
WM in ISLA	14
Motivation in ISLA	20
L2 exposure in ISLA	
IDs in WM, motivation, and L2 exposure as a complex dynamic system	
IDs in Russian ISLA: Future directions	45
Present study	47
CHAPTER 3: METHODOLOGY	49
Russian summer program: Curricular levels and distinct features	49
Russian summer program: Schedule and activities	57
Participants	
Placement and exit tests	60
WM tasks	67
Motivation surveys: L2MSS and General Motivation	69
Additional L2 exposure log	72
Post-program interviews	74
Procedure	76
Summary of Chapter 3	78
CHAPTER 4: PRE-STUDY PREPARATION	79
Internal Reviews	79
Pilot study #1: Test equating	79
Pilot study #2: Dissertation pilot study	81
Dissertation research modifications	83
Summary of Chapter 4	86
CHAPTER 5: DATA PROCESSING	
Data collection	

Data preparation	88
Initial data preparation	
Data cleaning, missing data, and outliers	91
Reliability and validity	93
Data input	96
Quantitative data	96
Qualitative data	
Processing and output	
Data storage	
Summary of Chapter 5	100
CHAPTER 6: MOTIVATION AND L2 EXPOSURE AS DYNAMIC CONSTRUCTS	101
Analysis overview of RQ 1	101
Motivation as a dynamic ID: Trends at four curricular levels combined	104
Motivation as a dynamic ID: Trends at different curricular levels	113
Motivation as a dynamic ID: Trends at the participant level	124
L2 Exposure as a dynamic ID: Trends at four curricular levels combined	131
L2 Exposure as a dynamic ID: Trends at different curricular levels	137
L2 Exposure as a dynamic ID: Trends at the participant level	142
Summary of Chapter 6	144
CHAPTER 7: THE ROLE OF IDs IN L2 RUSSIAN LEARNING	147
Analysis overview of RO 2	147
The role of WM in L2 Russian learning: Correlation analysis and interview insights	156
The role of motivation in L2 Russian learning: Correlation analysis and interview	
insights	
The role of additional L2 exposure in L2 Russian learning: Correlation analysis and	
interview insights	
The role of WM, motivation, and additional L2 exposure in L2 Russian learning: Reg	ression
analyses	163
Summary of Chapter 7	167
CHAPTER 8. GENERAL DISCUSSION AND CONCLUSION	172
Summary of the findings	172
General discussion	172
Discussion of the dynamics of motivation	175
Discussion of the dynamics of L2 exposure	182
Discussion of the role of IDs in L2 Russian learning	187
Theoretical implications: DST and future ISLA/ID research	193
Mixed-methods research and I CTI s	197
Ecological validity and ISLA research with I CTLs	200
Collecting data online during the COVID-19 pandemic	201
Practical implications	201
Pedagogical implications	203
Implications for the development of background questionnaires and placement	203
testing	206
woung	

Limitations and future directions	
Conclusion	210
APPENDICES	
Appendix A: L2MSS survey	
Appendix B: General Motivation survey	
Appendix C: Additional L2 exposure log	
Appendix D: Interview questions	
Appendix E: IRB approval	
REFERENCES	

LIST OF TABLES

Table 1. Number of participants at different curricular levels and phases of the study
Table 2. Key differences between the dissertation pilot study and dissertation research
Table 3. Outlier analysis (four curricular levels combined) 93
Table 4. Reliability analysis (motivation surveys; four curricular levels combined)
Table 5. Reliability analysis (dependent and independent variables; four curricular levelscombined)
Table 6. Main independent and dependent variables used in the analyses
Table 7. Descriptive statistics and 95% CI for the ID data (four curricular levels combined)103
Table 8. Descriptive statistics and 95% CI for the ID data (WM, Curricular Levels 1-4)111
Table 9. Descriptive statistics and 95% CI for the ID data (motivation, CurricularLevels 1-4)
Table 10. Descriptive statistics and 95% CI for the ID data (additional L2 exposure, Curricular Levels 1-4)
Table 11. Additional L2 Russian exposure in mean hours in descending order, Monday throughFriday, Phase 2 ($n = 34$) (four curricular levels combined)
Table 12. Additional L2 Russian exposure in mean hours in descending order, Monday throughFriday, Phase 3 ($n = 33$) (four curricular levels combined)
Table 13. Additional L2 Russian exposure in mean hours in descending order, Monday throughFriday, Phase 4 ($n = 29$) (four curricular levels combined)
Table 14. Additional L2 Russian exposure in mean hours in descending order, Monday throughFriday, Phase 5 ($n = 24$) (four curricular levels combined)
Table 15. Descriptive statistics and 95% CI for the entrance and exit testing scores
Table 16. Descriptive statistics and 95% CI for the learning gains
Table 17. Students' L2 speaking development during the summer program (number of studentsand % of all students in each level)150
Table 18. Students' L2 writing development during the summer program (number of students and% of all students in each level)150
Table 19. Relationships between the learning gains on the four tests
Table 20. Relationships between WM and the learning gains on the four tests

Table 21. Relationships between additional L2 exposure and the learning gains onthe four tests154
Table 22. Relationships between motivation and the learning gains on the four tests155
Table 23. Regression model (Phase 1): FIB entrance scores and WM Composite as predictors ofthe FIB exit scores (Curricular Levels 1-3)
Table 24. Regression model (Phase 1): MC entrance scores and WM Composite as predictors ofthe MC exit scores (Curricular Levels 1-3)164
Table 25. Regression model (Phase 5): MC entrance scores and General Motivation (Phase 5)as predictors of the MC exit scores (Curricular Levels 1-3)

LIST OF FIGURES

<i>Figure 1</i> . Interactive Compensatory Model of Learning with numbers indicating typical correlations between variables (Schraw et al., 2005)
Figure 2. Later development of Baddeley's WM model (Baddeley, 2010)
Figure 3. ACTFL inverted pyramid representing the functions of the major levels (actfl.org)62
Figure 4. Sample question from the FIB lexicogrammar quiz
Figure 5. Sample questions from the MC lexicogrammar quiz
Figure 6. Sequence of the Operation Span Task
Figure 7. Sequence of the Backward Digit Span Task
Figure 8. Example of L2MSS survey items in Qualtrics71
<i>Figure 9</i> . General Motivation survey in Qualtrics72
Figure 10. Additional L2 exposure log in Qualtrics74
Figure 11. Procedure
<i>Figure 12.</i> Motivation as a dynamic ID (General Motivation), four curricular levels combined, Phases 1-5
<i>Figure 13.</i> Motivation as a dynamic ID (L2MSS), four curricular levels combined, Phases 1 and 5107
<i>Figure 14</i> . Motivation as a dynamic ID (General Motivation), Curricular Level 1, Phases 1-5
<i>Figure 15</i> . Motivation as a dynamic ID (General Motivation), Curricular Level 2, Phases 1-5
<i>Figure 16</i> . Motivation as a dynamic ID (General Motivation), Curricular Level 3, Phases 1-5
Figure 17. Motivation as a dynamic ID (L2MSS), Curricular Level 1, Phases 1 and 5118
Figure 18. Motivation as a dynamic ID (L2MSS), Curricular Level 2, Phases 1 and 5119
Figure 19. Motivation as a dynamic ID (L2MSS), Curricular Level 3, Phases 1 and 5120
<i>Figure 20</i> . Motivation as a dynamic ID (General Motivation), Curricular Levels 1-3, Phases 1-5
<i>Figure 21</i> . Motivation as a dynamic ID (ideal L2 self), Curricular Levels 1-3, Phases 1 and 5

<i>Figure 22.</i> Motivation as a dynamic ID (ought-to L2 self), Curricular Levels 1-3, Phases 1 and 5
<i>Figure 23.</i> Motivation as a dynamic ID (anti-ought-to L2 self), Curricular Levels 1-3, Phases 1 and 5
<i>Figure 24</i> . Motivation as a dynamic ID (learning experience), Curricular Levels 2 and 3, Phases 1 and 5
<i>Figure 25.</i> Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), four curricular levels combined, Phases 2-5
<i>Figure 26.</i> Additional L2 Russian exposure in total hours in 12 categories, four curricular levels combined, Phases 2-5
<i>Figure 27.</i> Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 1, Phases 2-5
<i>Figure 28.</i> Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 2, Phases 2-5
<i>Figure 29.</i> Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 3, Phases 2-5
Figure 30. Additional L2 exposure as a dynamic ID, Curricular Levels 1-3, Phases 2-4143
<i>Figure 31.</i> WM Composite (Phase 1) as a predictor of the exit scores on the FIB and MC quizzes, Curricular Levels 1-3, error bars = 95% CI
<i>Figure 32.</i> General Motivation (Phase 5) as a predictor of the exit scores on the MC quiz, Curricular Levels 1-3, error bars = 95% CI
Figure 33. Simple mediation model with one mediator (Hayes, 2017)195
<i>Figure 34</i> . Simple moderation model with one moderator (Hayes, 2017)
Figure 35. Structural Equation Model with Parameter Estimates (Winke, 2013)196

KEY TO ABBREVIATIONS

BDS	Backward Digit Span (WM Task)
DST	Dynamic Systems Theory
FIB	Fill-in-the-Blank (Quiz)
ID	Individual Difference
ISLA	Instructed Second Language Acquisition
LCTL	Less Commonly Taught Language
L1	First Language
L2	Second Language
L2MSS	L2 Motivational Self System
MC	Multiple-Choice (Quiz)
OSpan	Operation Span (WM Task)
RQ	Research Question
SLA	Second Language Acquisition
XX / N /	XX7 1' X6

WM Working Memory

CHAPTER 1: INTRODUCTION

Situating the study

Why do L2 learners coming from similar backgrounds, even those from the same school cohorts, often achieve different results? A primary reason pertains to learners' IDs, a major topic in SLA and ISLA research (for review, see, e.g., Li et al., 2022; Loewen & Sato, 2017). IDs have been traditionally defined as attributes or personal characteristics that make learners unique and thus modify L2 acquisitional processes and the effectiveness of instruction (Dörnyei, 2013). In ISLA literature, linguists have used different classifications of IDs. For example, some of the important ID categories that have direct implications for L2 pedagogy described in the edited volume by Loewen and Sato (2017) included: social dimensions such as race, class, gender, sexuality, educational background, immigration status, and ethnicity (Duff, 2017); cognitive individual differences such as WM and L2 aptitude (Li, 2017); L2 motivation (Csizér, 2017); psychological dimensions of ISLA such as anxiety and other personality traits (Dewaele, 2017); teacher-related variables such as L2 instructors' first language (L1) background and teaching experience (Gurzynski-Weiss, 2017); students' ages and how they affect child SLA and adult SLA (Oliver et al., 2017); and heritage learner status (Montrul & Bowles, 2017). Another prominent ID category found in previous research (e.g., De Wilde et al., 2022; Pastushenkov, 2020; Ranta & Meckelborg, 2013) is the amount of L2 exposure from different sources (e.g., learning an L2 in formal and informal settings). This ID is particularly important from pedagogical perspectives, as individuals tend to have more control over the amount and type of L2 exposure that they are engaged in as opposed to their control over cognitive and psychological factors (Pastushenkov, 2020).

Rather than focusing on a single ID category, I strove to apply "a dynamic system lens to conceptualizing the interconnectedness of learner resources in instructed L2 development over time at varying proficiency" (Serafini, 2017, p. 370). Guided by DST (e.g., Ellis, 2007; Larsen-Freeman, 2014) and multidimensional perspectives on individual differences in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013), the present study focuses on three conceptually different categories of learners' IDs in adult ISLA: WM (a cognitive ID), motivation (referred to as an affective or a psychosocial ID), and differences in L2 exposure (i.e., L2 usage that includes linguistic input, output, and interaction, see Ranta & Meckelborg, 2013). Despite the previous propositions regarding IDs being interrelated (e.g., de Bot et al., 2007; Dörnyei, 2009a; Pawlak, 2012; Sparks et al., 2012; Sun, 2019), including WM and motivation (Serafini, 2017; Winke, 2013) and WM and L2 exposure (Denhovska et al., 2016; Indrarathne, & Kormos, 2018), the longitudinal development of these IDs (particularly of L2 motivation and L2 exposure) and their effects on learning gains at different stages of L2 development have not yet been investigated. This study aims to address this gap and explore how L2 learners' motivation and the amount of L2 exposure change over time and how WM, motivation, and the amount of L2 exposure affect learning gains in lexicogrammar, speaking, and writing. These three ID categories have been considered as key predictors of successful L2 development (WM and motivation, e.g., Serafini, 2017; L2 exposure, e.g., Muñoz, 2014; Ranta & Meckelborg, 2013) and thus an investigation of these constructs in one study can help better understand the role of IDs in instructed L2 learning. This study strives to reconsider problematic dichotomies (specifically the cognitive/affective dichotomy, see Dörnyei, 2010) and promote research guided by multidimensional perspectives on L2 learners' IDs (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013).

This study considers different methodological questions, propositions, and future directions outlined in previous SLA and ISLA literature on the topic of IDs. Considering the principles of DST (e.g., Larsen-Freeman, 2014), instructed L2 learning is viewed as a longitudinal process that involves different contexts and pedagogical interventions that can influence L2 learners' IDs (Pawlak, 2012). In line with previous ISLA research, including the empirical study by Serafini (2017), which was also guided by DST, motivation (Boo et al., 2015) and L2 exposure (Ranta & Meckelborg, 2013) have been considered as dynamic constructs, whereas WM as a cognitive factor was assumed to be more stable than fluctuating. In addition to these methodological considerations, the study addresses the need for cross-sectional ID research at multiple initial proficiency levels, as previous studies have shown that ID effects may vary at different stages of L2 development (WM and motivation, Serafini, 2017; WM, Serafini & Sanz, 2016; L2 exposure, Pastushenkov, 2020). Another important direction for ISLA research that this study aims to address is that it is important to investigate longitudinal effects of both linguistic (e.g., L2 exposure) and non-linguistic IDs (e.g., WM and motivation) at higher proficiency levels (e.g., Sagarra, 2017; Winke, 2013). Considering these propositions and the principles of DST, the present study includes multiple data points and different curricular levels (including students at higher proficiency levels) of an eight-week Russian summer immersion program.

Need for the study

Despite being more commonly taught and researched than many other LCTLs, Russian, with the exceptions of a few mostly quantitative studies (e.g., Denhovska & Serratrice, 2017; Denhovska et al., 2016; Kisselev et al., forthcoming) and theoretical work (e.g., Kisselev et al., 2020), remains underrepresented in ISLA literature. Even less is known about Russian learners' experiences (e.g., Klimanova & Dembovskaya, 2013; Merrill, 2020; Zaykovskaya et al., 2017)

and their motivation to study the language from qualitative and mixed-methods perspectives (Pastushenkov & McIntyre, 2020). In Russian ISLA research, studies have often been laboratory-based (e.g., P. Brooks & Kempe, 2013) with participants having no prior exposure to Slavic languages, thus putting the ecological validity of the studies in question when it comes to the effects of L2 instruction and relevance for ISLA. This limitation was also pointed out by Denhovska and colleagues who emphasized the need to test their findings longitudinally in ecologically valid classrooms settings (e.g., Denhovska et al., 2016). By conducting this study at an eight-week Russian summer immersion program, I aim to address the call for more ecologically valid SLA research (Godfroid, 2016), particularly in ISLA studies with LCTLs (Pastushenkov & McIntyre, 2020; Thompson, 2017), and the need to build stronger connections between L2 research and pedagogy (see, e.g., Sato et al., 2021). By adopting a mixed-methods approach with multiple data points, the study's goal was to obtain a more fine-grained view of IDs in instructed L2 learning, as well as to provide students' perspectives on IDs in ISLA. The project's goals are: (a) to better understand how conceptually different IDs (motivation and L2) exposure) change over time and how IDs (WM, motivation, and L2 exposure) affect longitudinal instructed L2 learning of different skills; and (b) to build an empirical foundation for the development of new background questionnaires, student placement policies, and instructional approaches and surveys that consider L2 learners' IDs and combine learner psychology and ISLA perspectives (see Sato & Csizér, 2021).

Significance of the study

To the best of my knowledge, this is the first longitudinal mixed-methods ISLA study focusing on cognitive, affective (or psychosocial), and exposure-related IDs with at a wide range of initial proficiency levels, study phases, and language skills. Morphologically rich languages

such as Russian require extended learning periods for most L1 speakers of English (Looney & Lusin, 2019), with various factors affecting the process of instructed L2 learning (e.g., Serafini, 2017; Winke, 2013). Therefore, a longitudinal perspective is vital in a Russian ISLA study. Even though the data were collected at an eight-week program, which is a considerably shorter time frame than those of some other longitudinal semester-long or year-long ID research (e.g., Sagarra, 2017), this study's context is an intensive immersion program that can yield learning gains comparable to one year of university instruction (see, e.g., Merrill et al., 2021; Rifkin, 2005). This study was also cross-sectional as it looked at several proficiency levels simultaneously. Collecting data in this Russian summer program was a unique and valuable opportunity that will help promote the generalizability of Russian ISLA research and its practical implications by extending previous laboratory-based research to an authentic setting (e.g., Denhovska, et al. 2016) and thus help promote the research-practice dial (Sato et al., 2021). Practical implications, including pedagogical considerations and implications for the development of placement tests and background questionnaires, will be discussed in this dissertation.

Despite the clear need to investigate the SLA/ISLA of LCTLs and critical languages (e.g., Kim, 2017; Pastushenkov & McIntyre, 2020; Thompson 2017), Russian remains understudied within the field. This project has the potential to spark an interest in the ISLA of LCTLs, including Russian. Adopting a dynamic systems lens and collecting both quantitative and qualitative data in this program can help obtain a deep understanding of the development of IDs and their role in instructed L2 Russian learning. Theoretical and methodological considerations regarding the application of DST and the mixed-methods methodology in ISLA research, including ID studies with LCTLs, will also be discussed in this dissertation.

Even though the sample size in the present dissertation is relatively low and issue of participant attrition was evident, particularly in later phases of the study and higher proficiency levels, this study is still an important step forward compared to many other previous studies with Russian and other LCTLs. The 2021 summer program brought 111 students, 52 of whom participated in the study.

According to the 2016 report of the Modern Language Association, there were more than 20,000 students enrolled in Russian language and culture programs in the United States (for an analysis of the enrollment in the Russian language programs and other issues, see Dengub et al., 2020). The results of this study will speak directly to learners and teachers of Russian, and possibly other languages. The project has the potential to illustrate the educational benefits of this immersion program and their pedagogy and has a number of other important theoretical and practical implications. Authentic ISLA data will be of interest to university administrators and teachers who can use this information to enhance language learning experience of their students.

Organization of the dissertation chapters

This dissertation includes eight chapters. Chapter 1 (Introduction) focuses on situating the study and explains the need for this research and its significance. Chapter 2 (Literature Review) is divided into several sections, focusing on L2 learners' IDs. This chapter starts with the description of DST and multidimensional perspectives on L2 learners' IDs, the overarching theory that guides the entirety of this study, including its procedure and analyses. The DST section is followed by the review of previous ISLA research on the topics of WM, motivation, L2 exposure, and the relationships between these ID categories. Chapter 2 also includes a section on ID research in the area of Russian ISLA. The chapter concludes with the present study section that includes two research questions (RQs). Chapter 3 (Methodology) starts with a detailed

explanation of the context of the study (the eight-week Russian summer immersion program), including the description of the program's curricular levels, distinct features, the students' schedule, and activities. This introductory part is followed by the description of the study's materials (the program's placement and exit test; WM tests; the L2MSS and General Motivation surveys; additional L2 exposure log; and post-program interviews). In Chapter 3, I also talk about the study's procedure and explain the analyses of the quantitative and qualitative data. Chapter 3 concludes with a summary. Chapter 4 (Pre-Study Preparation) includes the description of the internal reviews conducted at my home institution (Michigan State University) and in the summer program, as well as the description of the two pilot studies (test equating and dissertation pilot study) and the modifications in the dissertation study's methodology that I made upon completion of these pilot studies. Chapter 4 concludes with a summary. Chapter 5 (Data Processing) includes the description of several steps that I took when processing the study's data: data preparation (initial data preparation; data cleaning, missing data, and outliers; reliability and validity); data input (quantitative and qualitative data); processing and output; and data storage. Chapter 5 also concludes with a summary.

Chapter 6 (Motivation and L2 Exposure as Dynamic Constructs) and Chapter 7 (The Role of IDs in L2 Russian Learning) correspond to RQs 1 (parts A and B) and 2 of this study, respectively. These chapters include the analyses of quantitative and qualitative data at different curricular levels. Each chapter concludes with a summary. Chapter 7 (General Discussion and Conclusion) includes the summary of the findings based on RQ 1 (Chapter 6) and RQ 2 (Chapter 7). The summary is followed by the discussion of the study's theoretical and practical implications, the section about limitations, future directions, and the conclusion.

CHAPTER 2: LITERATURE REVIEW

Dynamic systems and multidimensional perspectives on IDs

The study is guided by DST (e.g., de Bot et al., 2007; Ellis, 2007; Hiver, 2015; Larsen-Freeman, 1997, 2012, 2014; Papi & Hiver, 2020; Serafini, 2017; van Geert, 1994; 1995; Waninge et al., 2014) and multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013). Originating in mathematics and physics, DST implies a continuous interaction of cognitive, social, and environmental factors in the acquisition process (de Bot et al., 2007; Ellis, 2007). In line with this theory, a growing number of L2 researchers adopted a broader, holistic approach that viewed "language, cognition, and language learners as dynamic complex, adaptive systems within which all factors interact and affect one another" (Serafini, 2017, p. 370). According to DST, a system (i.e., a set of variables such as IDs) is always nested within another system, and its development depends on the system's initial conditions, also known as the 'butterfly effect,' a concept originally proposed by the mathematician and meteorologist Edward Norton Lorenz that considers the huge impact of small local effects on global weather (de Bot et al., 2007). In line with these propositions, researchers have argued that L2 development shared similarities with complex ecosystems (e.g., forest ecosystems) in the natural world (e.g., Ellis, 2007; Larsen-Freeman, 1997). Individual factors within this ecosystem such as IDs mutually affect each other over time in a non-linear fashion (van Geert, 1994) and engage in supportive, competitive, conditional, or compensatory relationships (de Bot et al., 2007). For example, learners' effort can help compensate for the lack of time, or affective IDs such as motivation can help compensate for the limited input (a constituent of L2 exposure) from the learning environment (de Bot et al., 2007, p. 12).

A complex dynamic system can display both erratic and stable behaviors; stability is attributed to coming to so-called preferred "attractor" states (Serafini, 2017) or patterned outcomes (Hiver, 2015), resulting from "the propensity to self-organize" (Larsen–Freeman, 2012, p. 209) and making a system more predictable and thus more researchable (Dörnyei, 2010). An example of an attractor state in motivation would be a consistently positive value associated by L2 learners with living and studying abroad (see Papi & Hiver, 2020). Such variations in L2 development as evidenced by periods of erratic behaviors followed by stable patterns are directly linked to the context that pushes or pulls the system toward or away from a certain state (Waninge et al., 2014). Contexts can be characterized by cultural (e.g., the role of a teacher and a student in the target language), social (e.g., relationships between peers in a group), or other factors (Serafini, 2017).

Despite the propositions regarding the interconnectedness of IDs and hypothesized "selfreinforcing loops between the cognitive, motivational, social, and experiential aspects of L2 use" (Segalowitz & Trofimovich, 2012, p. 187), as Serafini (2017) and other proponents of DST pointed out, mainstream approaches have remained limited when it comes to conceptualizing IDs. Previous research has often assumed that IDs were "modular (i.e., discrete, separate, isolated), stable, and context independent traits" (Serafini, 2017, p. 370). In line with various SLA scholars (e.g., Dörnyei, 2010; Dörnyei & Ryan, 2015; Larsen-Freeman, 2014; Segalowitz & Trofimovich, 2012; Toth & Davin, 2016), Serafini supported the need to re-evaluate problematic dichotomies such as the cognitive/affective dichotomy, as well as to go beyond measuring IDs at one point and adding a longitudinal, dynamic lens to future research. Proponents of DST as a theory of SLA, as well as other L2 researchers, emphasized the need to further investigate dynamic interactions between different categories of IDs and their longitudinal development in

instructed L2 learning (e.g., Csizér, 2017; Dörnyei, 2009b; Pawlak, 2012; Sparks et al., 2012). Such propositions have been realized in studies at the intersection of cognitive and affective IDs (e.g., Serafini, 2017; Winke, 2013), as well as research focusing on cognitive and exposurerelated factors such as the amount of L2 input (e.g., Denhovska et al., 2016). L2 researchers have also pointed out that broader, holistic cognitive/psychosocial (Serafini, 2017) and sociocognitive perspectives (Toth & Davin, 2016) can help inform L2 pedagogy and thus would benefit ISLA as a field that aims to bridge the divide between L2 research and pedagogy (see, e.g., Sato et al., 2021).

To prepare an empirical ISLA study guided by DST, it is important to consider various practical issues and constraints. Hiver and Al-Hoorie (2016) outlined practical suggestions for L2 research guided by complexity/DST framework. In line with previous research under this framework (specifically the study by Spoelman & Verspoor, 2010), Hiver and Al-Hoorie came up with a "dynamic ensemble" (i.e., a practical catalogue or a blueprint for L2 researcher) that includes different practical considerations and questions under the four large categories¹ (p. 744): operational (e.g., "What is the complex system under investigation?"), contextual (e.g., "What are the contextual factors that are part of the environmental frame of reference for the system, its dynamic actions, and its patterned outcomes?"), macro-system (e.g., "What general principles of change exist for this system?"), and micro-structure (e.g., "What are the parts that make up the system under investigation?"). In light of these considerations and "the untapped potential" of the complexity/DST framework, Hiver and Al-Hoorie outlined key objectives for applied linguists conducting empirical research in this area, including the need to:

¹ These considerations guided this study's methodology, pre-study preparation, and data processing described in Chapters 3, 4, and 5.

(a) represent and understand specific complex systems at various scales of description;

(b) identify and understand the dynamic patterns of change, emergent system outcomes, and behavior in the environment;

(c) trace, understand, and, where possible, model the complex mechanisms and processes by which these patterns arise; and

(d) capture, understand, and apply the relevant parameters for influencing the behavior of the systems (p. 752).

The view of IDs as a complex, dynamic system with interconnected elements aligns with multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013). For example, Sun and Zhang (2020) concurrently investigated the speaking of advanced learners of L2 Chinese from cognitive, affective, and socio-cultural perspectives. The researchers argued that L2 learners' speech performance and production were affected by these three categories (or dimensions of IDs). However, as it was pointed out by Sun and Zhang (2020) and other researchers (e.g., Dörnyei, 2010; Segalowitz & Trofimovich, 2012), previous studies have often focused on one dimension of IDs rather than investigating different variables concurrently. The primary goals of Sun and Zhang were to address this gap and establish a comprehensive view of how multilinguals perceived their L2 Chinese speech using qualitative inquiry in the form of focus groups and semi-structured interviews. The researchers argued that their participants' production and performance of L2 Chinese speech may have been "the result of the synergistic effects of the cognitive (e.g., age of acquisition, cognitive fluency, learning styles, and speaking strategies), affective (e.g., motivation, anxiety, speaking self-efficacy, and WTC [willingness to communicate]), and socio-cultural (e.g., attitudes toward L2 Chinese culture, community, and classes) dimensions of L2 Chinese learning" (p. 11). Remaining

consistent with proponents of DST (e.g., de Bot et al., 2007; Ellis, 2007; Larsen-Freeman, 2014; Papi & Hiver, 2020; Waninge et al., 2014), these findings supported the need to consider conceptually different ID dimensions in ISLA research, including in qualitative ID studies.

Compensatory relationships between IDs, a central idea of DST, have also been discussed by various researchers and theorists working in the area of L2 reading (e.g., Bernhardt, 2005; Guo, 2018; McNeil, 2012). Many of these propositions were influenced by the idea of compensatory notion where readers lacking knowledge in one area switch to other input sources to compensate for such insufficiencies (Stanovich, 1984). For example, Bernhardt's (2005) model "tries to model how knowledge sources assist or take over for other knowledge sources that are inadequate or nonexistent—i.e., what they use to compensate for deficiencies" (p. 140). The three-dimensional model includes L1 literacy, L2 knowledge, and "unexplained variance" (domain-specific knowledge, motivation, etc.). McNeil (2012) extended on Bernhardt's (2005) ideas and proposed a model of reading that includes L2 language knowledge, L1 reading ability, strategic knowledge, and background knowledge at two proficiency levels. These theoretical propositions were empirically tested by Guo (2018), who found evidence supporting the facilitative role of L1 reading ability and L2 language proficiency in L2 reading.

Complex relationships between learners' IDs have also been discussed outside the domains of SLA and ISLA (e.g., D. Brooks & Shell, 2006; Schraw et al., 2005; Verhaeghen et al., 2012). Proposed by Schraw et al. (2005), the Interactive Compensatory Model of Learning (ICML) includes the following components: cognitive abilities, a knowledge base, learning strategies, metacognition, and motivation² (see Figure 1). Similarly to DST, ICML emphasizes

 $^{^{2}}$ A discussion of learning strategies and metacognition goes beyond the scope of this dissertation; these IDs, however, are considered as important future directions.

the interconnectedness of variables in its effect on acquisitional processes. According to this model, learners can compensate for weaknesses in one of the areas by using their knowledge and strengths in others (Schraw et al., 2005). D. Brooks and Shell (2006) suggested that teachers can adopt this model in their classes (e.g., by explicitly teaching how to apply learning strategies to compensate for the lack of metacognitive skills). For example, the interplay between WM and motivation has been discussed in the context of chemistry teaching (e.g., D. Brooks & Shell, 2006; Schraw et al., 2005). Based on ICML, "motivation is the process by which we consciously or unconsciously allocate WM resources" (D. Brooks & Shell, 2006, p. 24). However, as D. Brooks and Shell (2006) pointed out, an interpretation of motivation explicitly in terms of WM has not yet been explored.



Figure 1. Interactive Compensatory Model of Learning with numbers indicating typical correlations between variables (Schraw et al., 2005).

WM in ISLA

As Li (2017) pointed out, the considerable interest of L2 researchers in cognitive IDs (particularly in L2 aptitude and WM) was motivated by their explanatory power with regards to differences in L2 learning outcomes which made this ID category an important area of inquiry in ISLA research (e.g., Ahmadian, 2020; Duman et al., 2021; Lee & Revesz, 2021; Malone, 2018; Nielson & DeKeyser, 2019; Ruiz et al., 2021; Sagarra, 2019). Pedagogically speaking, understanding WM, one of the ID categories under investigation in this dissertation, can help practitioners "tailor instruction to achieve maximal instructional effects" (Li, 2017, p. 396). Before discussing the pedagogical implications of WM research, it is important to discuss how this ID has been conceptualized and measured in ISLA literature and what were some gaps found in previous WM research in instructed L2 learning settings, particularly in longitudinal studies similar to this dissertation research.

Miyake and Friedman (1998) defined WM as "a computational arena or workspace, fueled by flexibly deployable, limited cognitive resources, or activation that support both the execution of various symbolic computations and the maintenance of intermediate products generated by these computations" (p. 341). Although various researchers have proposed their theoretical conceptualizations of WM (e.g., Miyake & Friedman, 1998), as Serafini and Sanz (2016) and other L2 researchers (e.g., Li, 2017) pointed out, Baddeley's domain-specific, multilevel resource model (see Figure 2) remains dominant within SLA and ISLA research (e.g., Baddeley, 2010). This "nonunitary" model includes the central executive component (an attentional control system) supported by two short-term storage systems connected by the episodic buffer: the visuo-spatial sketchpad for visual material and the phonological loop for verbal-acoustic material (Baddeley, 2010).



Figure 2. Later development of Baddeley's WM model (Baddeley, 2010).

To explore the role of the different components of WM, researchers within the domains of cognitive psychology and ISLA have often used WM span tasks (for review, see Conway et al., 2005). Span tasks can be verbal or non-verbal and vary in complexity; however, only complex span tasks can push learners to their limits and thus capture the limits of WM capacity (Serafini & Sanz, 2016). For example, the operation span (OSpan) task developed by Turner and Engle (1989) and later updated to an automated version by Unsworth et al. (2005) involves solving arithmetic equations (e.g., "Is $(1 + 3) \ge 2 = 8$ true or false?") while maintaining words or letters in memory during reading or listening and then recalling them. Different versions of OSpan have been widely used in SLA and ISLA research (e.g., Faretta-Stutenberg & Morgan-Short, 2018; Serafini & Sanz, 2016). In addition to OSpan, researchers have also commonly used Digit Span tasks (Serafini & Sanz, 2016; van den Noort et al., 2006). In these tasks, participants hear or see a series of digits (one by one; usually up to 9 digits in a series) that they need to repeat either in the order of presentation or backward. Digit Span tasks measure storage components of WM, whereas complex span tasks such as OSpan tap into the central executive component that coordinates storage and processing (Hale et al., 2011). WM tasks are often administered in the participant's L1 to control for a potential confounding variable with their verbal processing (Serafini & Sanz, 2016; van den Noort et al., 2006). Both OSpan and Digit Span tasks were reported to have high reliability in research with college-age participant populations (e.g., Conway et al., 2005).

Over the years, various researchers have found empirical evidence supporting the facilitative role of WM capacity in SLA and ISLA (e.g., Erçetin & Alptekin, 2013; Indrarathne & Kormos, 2018; Linck & Weiss, 2015, Mackey et al., 2010). The meta-analysis by Linck et al. (2014) focusing on studies in instructed settings showed that a higher WM capacity positively affected L2 processing and proficiency outcomes (an estimated population effect size (ρ) = .255), with larger effects found for the executive component of WM (as opposed to the storage components). Research has shown that WM capacity not only positively affected learning under explicit and implicit learning conditions (Indrarathne & Kormos, 2018), but also influenced both implicit and explicit L2 knowledge (Erçetin & Alptekin, 2013). Moreover, researchers have found strong positive associations between a higher WM capacity and the production of modified output (Mackey et al., 2010).

Despite these findings, previous research has revealed different and even contradictory findings on how WM affected L2 development in the long term (for review, see Sagarra, 2017). Some studies have shown positive associations between a higher WM capacity and L2 long-term learning of different skills (e.g., grammar, see Linck & Weiss, 2011, 2015; grammar and reading,

see Sagarra, 2017; reading, listening, speaking, vocabulary, and grammar, see Kormos & Sáfár, 2008), whereas others have failed to support the facilitative role of WM (e.g., grammar, see Frost et al, 2013; grammar and vocabulary, see Grey et al., 2015). For example, Frost et al. (2013) did not find any positive associations between a higher WM capacity and L2 morphosyntactic processes in intermediate and advanced learners of Hebrew. Similarly, Grey et al. (2015) suggested that morphosyntactic and lexical development were not affected by the variation in WM and phonological WM over the course of a 5-week study abroad program in Spain. In contrast with these findings, Sagarra (2017) found evidence that WM played a facilitative role in grammar and reading development in beginning classroom learners. Interestingly, the participants in the studies by Grey et al. (2015) and Sagarra (2017) were native speakers of English who were studying Spanish. Even though the study by Sagarra (2017) was conducted over a longer time period in a different setting, there were other factors that may have influenced the contradictory results (no long-term WM effects vs. long-term WM effects).

Sagarra (2017) suggested the key factors affecting the results were the cognitive demands of WM tests and language tasks used by the researchers, as well as the participants' L2 proficiency. Cognitively demanding WM tests yielded long-term WM effects (e.g., Backward Digit Span task, see Kormos & Sáfár, 2008; Waters and Caplan's test, see Sagarra, 2017; Operation and Digit Span tasks, see Serafini & Sanz, 2016). Researchers have also found stronger associations between WM and performance on more difficult language tasks such as fill-in-the-gaps as opposed to easier ones such as picture-matching while listening (see Santamaria & Sunderman, 2015). Finally, as Sagarra (2017) pointed out, more robust WM effects have been found in beginner learners (e.g., Linck & Weiss, 2011, 2015; Serafini & Sanz, 2016) than in more advanced students (e.g., Frost et al., 2013; Grey, Cox, Serafini, & Sanz,

2015). In addition to the limitations outlined by Sagarra (2017), it is important to note that previous research has predominantly collected WM data at one point (e.g., Sagarra, 2017; Winke, 2013), as cognitive IDs such as WM and L2 aptitude have been generally considered as more stable constructs than other IDs, including affective motivational variables (e.g., Serafini, 2017). It is important to note, however, that Jackson (2020) suggested that "the construct of WM is inherently dynamic and complex, being a multi-faceted system that links storage and processing components, influences outcomes, and is shaped by time and experience" (p. 103). Considering Jackson's (2020) propositions, as well as the evidence of changes in WM capacity as a result of WM training (see Hayashi, 2019) and cognitive decline as a result of aging and underlying neurologic causes (e.g., Birdsong, 2018), future research, particularly ISLA studies conducted over the course of several semesters, may also consider adding additional points when WM data will be collected, particularly in studies conducted over the course of several years.

Considering the increased interest in WM and its power with regards to explaining differences in L2 learning outcomes, as well as and the proposition regarding the need to control for WM in ISLA research (e.g., Sagarra, 2019), this ID has been more commonly investigated in recent pedagogically oriented ISLA studies (e.g., Ahmadian 2020; Duman et al., 2021; Malone, 2018; Nielson & DeKeyser, 2019; Ruiz et al., 2021). For example, Ahmadian (2020) found evidence that WM was strongly and positively correlated with English language learners' gains on a discourse completion test and a pragmatics comprehension questionnaire on both the immediate and delayed post-tests under implicit learning conditions; however, no associations between a higher WM capacity and the students' learning gains were found in the explicit and control groups. Remaining consistent with the pedagogical suggestions of Li (2017), this finding demonstrated that the effectiveness of explicit instruction tended to be less constrained by L2

learners' WM capacity. Further emphasizing the mediating role of instruction, Nielson and DeKeyser (2019) found evidence that different task conditions affected L2 learners with highand low WM capacities differently and how they were able to comply with story-telling instructions that required them to focus on grammatical form. With regards to vocabulary learning under incidental conditions, Malone (2018) found evidence that WM scores correlated with vocabulary outcomes and emphasized that "implementing audio support for reading in the classroom need not exclude learners with lower WM ability" (p. 672). In line with the findings of the study by Malone, Ruiz et al. (2021) also found evidence that WM was associated with vocabulary L2 learning; however, this association depended on the instructional context and WM was predictive of L2 learning gains only in the form-focused condition. The role of WM in L2 learning of vocabulary and grammar, particularly the fact that WM capacity is limited, has also been considered in previous literature about different memorization strategies (e.g., multimediaassisted self-learning materials and digital flashcards, Zhu et al., 2012; Li & Tong, 2019; color coding, Ibarra Santacruz, & Martínez Ortega, 2018; mnemonic strategies, Di Santo et al., 2020). For example, due to the fact that WM is limited, using both on-screen text and animation simultaneously in digital flashcards may result in split attention effects and negatively affect learning (Zhu et al., 2012).

In recent years, ISLA research on the role of WM in different learning contexts has also adopted more advanced research methodologies. For example, in their eye-tracking study of the effects of captions, Lee and Revesz (2021) found evidence that different versions of tasks measuring L2 learners' visual short-term memory were associated with oral production gains in different groups. In line with the meta-analysis by Linck et al. (2014), the researchers pointed out that different components of WM may have played different roles under different learning conditions, further emphasizing the role of context as important variable to consider in ISLA research of WM (e.g., Serafini, 2017). It is also important to note that the role of WM in instructed L2 learning may also vary depending on the language skill under investigation. For example, the study of the effects of listening strategy instruction in relation to cognitive IDs (WM and L2 aptitude), Duman et al. (2021) did not find evidence of supporting role of a higher WM capacity or L2 aptitude in listening comprehension scores, even though the researchers pointed out the importance of the task type in WM research, as previous research has shown that a higher WM capacity played a stronger role in more complex tasks (see, e.g., Sagarra, 2017, 2019). Remaining consistent with DST principles (e.g., Hiver & Al-Hoorie, 2016; Waninge et al., 2014), L2 researchers tended to emphasize the role of context as an important variable to consider in ISLA studies of WM (e.g., type of instruction, Ahmadian 2020; type of task, Nielson & DeKeyser, 2019; and task difficulty, Duman et al., 2021).

In addition to addressing the gaps in previous longitudinal WM research, including the pedagogically oriented ISLA studies discussed in this section, it is important to note that the vast majority of previous WM studies and cognitive SLA and ISLA research in general focused exclusively on cognitive mechanisms of L2 learning (Serafini, 2017), even though "cognitive abilities comprise only part of any picture" (Mackey & Sachs, 2012) and "individual differences in mental functions typically involve a blended operation of cognitive, affective and motivational components" (Dörnyei, 2010, p. 234).

Motivation in ISLA

L2 motivation, which is another ID category under investigation in this dissertation research, has been considered one of the most important variables in recent ISLA literature (e.g., Fathi et al., 2019; Han & Hiver, 2018; Nagle, 2018; Papi & Hiver, 2020; Pfenninger &

Singleton, 2016; Sato & Csizér, 2021; Sato & Lara, 2019). This multi-dimensional ID has been viewed as a vehicle for L2 learning (Sato & Loewen, 2019a) that encapsulates the effort the learners are willing to invest in their L2 development (Csizér, 2017, Dörnyei, 2009a, 2009b, 2010). Before discussing pedagogical implications and future theoretical and methodological directions for L2 motivation research, it is important to examine how this important ID has been conceptualized and measured in previous literature.

The way motivation has been viewed in L2 theory and research goes back to Gardner and colleagues' socio-educational model that considered this ID as synonymous of effort and enthusiasm; it included three main components: integrativeness, attitudes toward the learning situation, and motivation (e.g., Gardner et al., 2004; Gardner & Lambert, 1972; Masgoret & Gardner, 2003). Integrativeness is a socio-psychological concept that represents to what extent L2 learners identify themselves and want to be associated with the L2 communities, which is also associated with L2 learners' intrinsic motivation. The attitudes component of the socio-educational model encapsulates learners' views of their L2 teachers and courses, which is associated with certain levels of anxiety. Finally, motivation, another component of Gardner and colleague's socio-educational model, is conceptualized as a combination of effort, desire, and attitude toward the L2 learning process.

The model proposed by Gardner and colleagues has led to the development of the standardized Attitudes & Motivation Test Battery (AMTB), which includes a series of questions that L2 learners self-report on a Likert-type scale in the three categories. For example, the AMTB questionnaire used by Serafini (2017), which is one of the studies that guided this dissertation research both theoretically and methodologically, included 78 items on a 7-point Likert scale into three sub-scales (integrativeness, attitudes, and motivation). These subscales

were reported to have high reliability based on Cronbach's alpha (see the meta-analysis by Masgoret & Gardner, 2003). Examples of an AMTB survey item would be, "It worries me that other students in my class seem to speak Spanish better than I do." or, "My Spanish Teacher: Efficient ____: ___: ___: ___: ___: ___: Inefficient) (Serafini, 2017, p. 377).

From Gardner and colleagues' conceptualization of L2 motivation as a multi-layered construct in their socio-educational model and socio-psychological perspectives, the field of SLA showed "a greater interest in concepts being developed in contemporary cognitive and educational psychology" (Boo et al., 2015, p. 146) with L2 motivation now being more commonly viewed through the prism socio-dynamic models (Serafini, 2017). In these models, L2 motivation has also been considered as a multi-dimensional ID and conceptualized in terms more than one concept (for review, see Boo et al., 2015), including selves or 'future self-guides,' as associated with the Dörnyei's (2009a) L2 Motivational Self System (L2MSS). Dörnyei's framework has been widely used in previous ISLA research (e.g., Pastushenkov & McIntyre, 2020; Serafini, 2017; Thompson, 2017). The traditional L2MSS, which is rooted in the mainstream motivation research in psychology (Boo et al., 2015), includes three components: two self-visions (ideal L2 self and ought-to L2 self) and learning experience. Similarly to AMBT surveys, L2 learners self-report their motivation on different versions of L2MSS questionnaires using items on a Likert-type scale. For example, Serafini's (2017) L2 self motivation questionnaire included 30 items on a 7-point Likert scale into three categories (10 items in each category): the ideal L2 self (i.e., characteristics that an L2 learner would like to have); the oughtto L2 self (i.e., characteristics that L2 learners think they ought to have); and motivated learning behavior (i.e., the effort learners intend to put in L2 learning). According to Dörnyei (2010), the reliability coefficients for these three categories were high. An example of an L2MSS survey

item would be, "I can imagine myself living abroad and having a discussion in Spanish" (Serafini, 2017, p. 377).

A strong vision of the ideal self implies that L2 learners can visualize who they want to become, while the ought-to self encapsulates external pressures (e.g., from teachers, parents, and friends) that learners experience (Thompson, 2017). Ideal L2 self has often been associated with intrinsic motivation, which refers to "doing an activity for its inherent pleasure and enjoyment rather than external consequences," and ought-to L2 self has been associated with extrinsic motivation, which refers to "doing an activity for its instrumental value rather than for pleasure or enjoyment" (Lai, 2013, p. 94). If L2 proficiency is a part of an individual's vision of ideal L2 self or ought-to L2 self, this individual will be motivated to acquire the L2 because of the "psychological desire to reduce the discrepancy between current and future self states" (Ushioda & Dörnyei, 2012, pp. 400-401).

Research has shown that students' visions of their ideal L2 and ought-to L2 selves often varied when they started learning an L2, with these visions being aligned with learners' personal and professional goals (e.g., Pastushenkov & McIntyre, 2020). In previous literature, these differences at the onset of L2 learning have been referred to as initial motivational profiles (see, e.g., Han & Hiver, 2018). For example, mastering English as a L2 has often been viewed as an opportunity to find a job at a transnational company and thus motivated students to start learning this language (e.g., Sun, 2019). The need to learn an L2 to assist in future college or job applications or support for other benefits of L2 learning could also be imposed on learners by their teachers or parents and thus associated with the ought-to L2 self. For example, Yana, a graduate of a Russian summer immersion program and one of the two focal participants in the motivation case study by Pastushenkov and McIntyre (2020) said, "This is secondary but... and I
also talked about it with my parents... this [learning Russian in an immersion program] was going to be an opportunity that would show me at a younger age what college was going to be like" (p. 18).

This case study also showed that a strong vision of learners' ideal L2 self could be associated with L2 learners' confidence. Katia, the other focal participant in the case study and a graduate of the same immersion program, was confident in her ability to speak Russian from the very start of her studies, especially because she had a chance to visit Ukraine (the birthplace of her adoptive brother) prior to her studies in the Russian summer program. This finding was supported by both quantitative data from the L2MSS survey and the qualitative data from the interview. Katia's confidence was reflected in the strong vision of her ideal L2 that remained consistently high over the course of her Russian language studies: before the summer immersion, immediately after the summer program, and several years later when Katia lived in a Russianspeaking country, with the latter period being associated with the strongest vision of her ideal L2 self. For both Katia and Yana, their motivation to study Russian was a combination of personal and professional reasons; however, Yana's professional goals to become an aerospace engineer have overshadowed her personal motivations (e.g., her love for the works for Leo Tolstoy). It is also important to note that Katia's and Yana's visions of their ideal L2 self and ought-to L2 self fluctuated over time (even though these visions remained more stable for Katia), emphasizing the need to view and investigate L2 motivation as a dynamic construct (e.g., Gardner et al. 2004; Larsen-Freeman, 2014; Papi & Hiver, 2020; Waninge et al., 2014).

Despite the differences in their visions of ideal L2 self and ought-to L2 self, both Katia and Yana positively viewed their L2 Russian learning experience in the summer immersion program. Learning experience (i.e., the "Cinderella" of the L2MSS as Dörnyei called it) is the third component of the traditional Dörnyei (2009) framework. This dimension of L2MSS is viewed as a "perceived quality of the learners' engagement with various aspects of the language learning process" (Dörnyei, 2019, p. 26). The "Cinderella" metaphor is often justified; for example, some of the studies that adopted L2MSS as their motivational framework focused on students' selves and did not include "learning experience" as a variable (e.g., Serafini, 2017). As Dörnyei (2019) pointed out, there is also a need for a refinement and a reconceptualization of learning experience as a component of L2MSS. According to the researchers, future studies will need to consider such constituents of L2 learning experience as the role of school context, syllabus and teaching materials, learning tasks, and learners' relationships with their peers and teachers.

It is important to note that L2MSS has been primarily adopted in research with learners of English and there has been a "need to study the motivation to learn languages other than global English" (Boo et al., 2015, p. 156). The dominance of English and other more commonly taught languages such as Spanish in previous motivation literature may have created a bias in assessing L2 motivation (Lanvers, 2016) and alternative options need to be explored in the study of LCTLs such as Russian (Pastushenkov & McIntyre, 2020). Various L2 researchers have suggested adding a new component in the traditional L2MSS framework: "rebellious" (Lanvers, 2016) or anti-ought-to L2 self (Thompson, 2017, 2021). Strong anti-ought-to L2 self-visions imply that learners choose to study a L2 despite being discouraged by others (Pastushenkov & McIntyre, 2020). Thompson (2021) compared this type of self to children proving adults wrong by saying, "See? I CAN do it!"

Considering the important pedagogical implications of L2 motivation research, this ID has become an integral component of the pedagogically oriented domain of ISLA research (for

review, see Csizér, 2017). For example, research has shown that L2 learners' sensory/imagery capacity in terms of their vision of future selves can be intentionally harnessed (Dörnyei & Chan, 2013) or trainable using different visualization techniques and helping them develop multilingual identities (e.g., Dornyei & Kubanyiova, 2014; Henry, 2017; Sato & Csizér, 2021; Sato & Lara, 2019). These pedagogical interventions can be particularly helpful because previous studies that used L2MSS questionnaires have shown positive relationships between L2 learners' desired selves (specifically, the ideal L2 self) and their learning effort and achievement (Dörnyei & Chan, 2013). For example, intensive immersion programs that adopted communicative and taskbased approaches to L2 pedagogy and that strove to facilitate a learner-centered classroom have been shown to positively affect L2 learners' vision of their ideal L2 selves that were associated with high levels of achievement in the long-term (see Pastushenkov & McIntyre, 2020). As was pointed out by Wang and Eccles (2013), "teachers who balance structural support and student autonomy in an effort to maintain an engaging environment are more likely to encounter increased learning motivation in their students" (p. 20). As an important construct of learner psychology, motivation can have mediating effects on feedback effectiveness (Sato & Csizér, 2021) and plays an important role in shaping L2 learners' pronunciation (Nagle, 2018). Previous research has also shown that motivation can help predict L2 learning success for both early and late starters (Pfenninger & Singleton, 2016). However, despite the "unprecedented boom" in the study of L2 motivation (Boo et al., 2015, p. 145) and the fact that this area has become "one of the most vibrant fields of applied linguistics" (p. 418), Csizér (2017) pointed out there is a clear need for more empirical research in instructed L2 settings, as the distinction between L2 motivation in instructed and naturalistic settings has remained underexplored.

In terms of theoretical considerations, the dynamic systems approach has been recently introduced in the domain of L2 motivation theory and research (e.g., Larsen-Freeman, 2014; Papi & Hiver, 2020). The view of L2 motivation as a dynamic construct (e.g., Waninge et al., 2014) aligns with previous conceptualizations of this ID, including Gardner and colleagues' socio-educational model (e.g., Gardner et al. 2004) and Dörnyei's (2009a) L2MSS. Larsen-Freeman (2014) outlined ten lessons that L2 motivation researchers can learn from this theory: nothing in a complex system is fixed; time and space are at the foremost of the theory's agenda; the novel behavior of a system emerges through the interaction of its components; a complex system comprises interrelated factors; its movement through space and time is non-linear and difficult to predict; the system depends on its initial condition (the 'butterfly effect'); a complex system always remains open while interacting with the environment; a system is feedbackdependent; it is context-dependent; and it has a non-Gaussian 'heavy-tailed' distribution (relying only on an average behavior is not informative).

In their qualitative study of L2 motivation, Papi and Hiver (2020) pointed out that the complexity/DST framework "has broad potential to open the conversation up to wider theoretical perspectives in L2 motivation as it takes a bird's eye perspective of the phenomenon" (p. 226). Guided by DST and using retrospective narratives, the researchers explored interactions between different motivations, contextual factors, and subsequent trajectories of the motivation of six Iranian doctoral students learning English in the United States. In their Global Model of Motivation Mechanisms, Papi and Hiver (2020) used process tracing to identify the following changes in the participants' motivation trajectories that emerge from learners' internal and contextual factors: initial conditions, adaptive interactions, coordinative structure, perturbation, self-organized criticality, and attractor state (p. 217). Initial conditions referred to the

participants' English language learning in Iranian middle schools. An example of adaptive interactions would be the learners' realization of a central role of English in their lives (i.e., additional value associated with English language learning), which was also something realistic that the learners were able to control. An example of the manifestation of coordinative structure would be the decision of parents to have their children attend a private school in Iran, which was associated with both positive (i.e., increased value of English language learning) and negative influences (i.e., the lack of control over this decision). An example of perturbation would be the high-stakes exam Konkoor that "acted as a disturbing force with the potential to shift all learners out of stability and into a different direction of language development and action" (p. 217). As an example of self-organized criticality, Papi and Hiver (2020) talked about the influence of teachers, particularly during secondary school, that played a major role in shaping the participants' pathways. The attractor state (i.e., a pattern of stability), an important construct in DST literature (see, e.g., Hiver, 2015; Serafini, 2017), was the new state when the extreme value was associated with living and studying at a North American university. The study by Papi and Hiver (2020) highlighted the importance of external factors (e.g., parents and teachers), particularly at the early stages of L2 development. For example, one of the participants stopped learning English at a certain point because of the lack of adequate attention from his new teacher. Pedagogically speaking, the participants mentioned that their teachers' instruction style and their learning experience were the key factors that made the students enjoy the learning process and continue to learn English.

Another example of recent empirical ISLA research of L2 motivation guided by DST was the quantitative study by Serafini (2017).³ In addition to motivation, the researcher also investigated WM in the instructed longitudinal L2 learning of Spanish. Serafini aimed to reevaluate the cognitive/affective dichotomy and adopted DST as an alternative theoretical framework that took into consideration the interrelatedness of IDs and their temporal dynamics. The researcher operationalized L2 motivation using both the AMTB and L2MSS. In addition to their empirical validation, these two motivational frameworks aligned with the view of learners' motivation as a complex and dynamic system (Serafini, 2017). Theoretically speaking, the key to Serafini's (2017) study was the cognitive context (i.e., the intrinsic dynamics of the L2 learner). Using van Geert's (1995) precursor model, which identifies precursors, successors, and 'connected growers' as variables or subsystems, Serafini conceptualized internal L2 learner resources such as their WM capacity as connected 'growers.' These 'growers' involved cooperation between cognitive and motivational components, and their relationships throughout L2 development were considered mutually supportive, competitive, conditional, and compensatory (see, e.g., de Bot, 2008). More efficient memory was hypothesized to be associated with a more positive motivational orientation; less efficient memory was hypothesized to be compensated by strong motivation. Moreover, the nature of these relationships was expected to fluctuate over time (as measured by minutes, days, weeks, months, and years; see de Bot, 2015). Serafini (2017) investigated what insights a DST perspective can bring to bear on the patterns of relationships between WM and L2 motivation and to what extent these relationships fluctuated for increasing L2 proficiency. The results of the analysis provided some evidence that

³ Considering the theoretical and methodological importance of the study by Serafini (2017) for this dissertation research, more detail about the work by Serafini will be discussed in the "IDs in WM, Motivation, and L2 Exposure as a Complex Dynamic System" section.

the cognitive (WM) and psychosocial (motivational) subsystems did not "operate in isolation from one another, but rather form part of the interdependent structure underlying the learner's internal cognitive ecosystem" (Serafini, 2017, p. 382), remaining consistent with previous propositions in line with DST (e.g., van Geert, 1995).

Methodologically speaking, as opposed to the predominantly quantitative cognitive ISLA research, studies of L2 motivation have been quantitative (e.g., Serafini, 2017; Thompson, 2017), qualitative (e.g., Papi & Hiver, 2020; Thompson & Vásquez, 2015; Sun & Zhang, 2020) and mixed-methods (e.g., Fathi et al., 2019; Pastushenkov & McIntyre, 2020). Being considered a promising methodology in the field of L2 motivation (Boo et al., 2015), the mixed-methods approach helped Fathi et al. (2019) explain the dynamics of L2 writing as supported by blogs, while considering the role of writing motivation, self-efficacy, and self-regulation. In the case study by Pastushenkov and McIntyre (2020), which adopted a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018), the qualitative data from the interviews and artifacts (examples of students' writing) helped clarify some of the patterns identified in the quantitative data from the L2MSS questionnaire. However, as it was pointed out by Boo et al. (2015) with regards to the mixed-methods approach in L2 motivation research, "most of the current practice seems to involve rather superficial mixing of relatively independent qualitative and quantitative components within a study" (p. 156) and therefore, more robust mixed-methods studies of L2 motivation are needed in the field of ISLA.

The literature review regarding L2 motivation in ISLA research once again showed that there has been a clear divide between cognitively-oriented L2 researchers (for review, see Li, 2017) and proponents of psychosocial approaches rooted in social and educational psychology (for review, see Csizér, 2017). As Serafini (2017) pointed out, cognitively oriented research

focused on learner-internal cognitive factors and strove to explore causal relationships between cognitive factors and L2 development, whereas proponents of psychosocial approaches such as researchers working in the area of L2 motivation generally considered L2 learning as a unique process that is unlike other skills (i.e., they viewed L2 development as an endeavor that is constrained and facilitated by psychological, social, and cultural factors specific to different contexts). In addition to addressing the need to re-consider the cognitive/affective dichotomy, it is also important to investigate other IDs that engage in complex relationships with cognitive and affective variables. One such important ID category is L2 exposure.

L2 exposure in ISLA

The final ID category under investigation in the present study, L2 exposure, is conceptualized as a combination of linguistic input and output from different sources (Ranta & Meckelborg, 2013), including in the form of interaction with teachers and peers (for a review of interaction research in ISLA, see Loewen & Sato, 2018). Input, output, and interaction have been considered key elements of L2 development (e.g., Gass & Mackey, 2006; Long, 1996) and as their amounts vary by learner, L2 exposure has been viewed as an important ID (e.g., Berghoff, 2022; Gallo et al., 2021; Kim et al., 2022; Pastushenkov, 2020). The theoretical framework around interaction, a concept directly linked to input and output (Ranta & Meckelborg, 2013), falls between two camps: Cognitive-Interactionist (e.g., Gass & Mackey, 2006; Loewen & Isbell, 2017; Long, 1996; Pastushenkov et al., 2021, Philp et al., 2010) and Sociocultural (e.g., Sato & Ballinger, 2012; Swain & Lapkin, 1998, 2000). In line with the interaction hypothesis (Long, 1996), a cognitive-interactionist framework considers the following constructs as the key elements of interaction: input, negotiation, output, and noticing (Loewen & Sato, 2018). Sociocultural theory, on the other hand, views L2 development as a dynamic process where

knowledge is co-constructed through discourse between a novice and an expert (Swain & Lapkin, 1998, 2000). Drawing upon Vygotsky's sociocultural theory (1978), socialization theorists have argued that social interaction shapes learners' cognitive development (Biedroń & Pawlak, 2016a). On the cognitivist side, Ranta and Meckelborg (2013) also supported the need to consider sociopsychological and sociocultural factors affecting L2 learners. Despite the subtle differences between the frameworks, the key proposition remains: interaction is an indispensable component in L2 development and is shaped by such factors as the students' L1 status, the interlocutors' role (peer interaction vs. teacher interaction), participation structure, L2 proficiency, and a variety of cognitive and psychological IDs (Loewen & Sato, 2018).

Different L2 researchers have pointed out the need to quantify exposure (e.g., Freed et al., 2004; Hanzawa, 2021; Isabelli-García & Lacorte, 2016; Kim et al., 2022; Merrill, 2020; Muñoz, 2008, 2014; Ranta & Meckelborg, 2013; Saito, 2015). In line with scholars working with other IDs such as motivation (e.g., Csizér, 2017), Muñoz (2008) pointed out the importance of the distinction between L2 exposure in naturalistic and instructed settings. The researcher emphasized that input in naturalistic settings is "generally unlimited," while in instructed settings, "input is, by definition, limited and it is usually distributed in very small doses" (p. 590). It is important to note, however, that the variation in the amount of L2 input in naturalistic L2 environments may be context/learner dependent. For example, the study by Ranta and Meckelborg (2013) showed that living, studying, or working in a naturalistic setting did not guarantee high levels of interaction in the L2, particularly in informal settings. Considering that input in instructed settings (and sometimes in naturalistic settings) can be limited, it can also be quantified. In a notable study that attempted to quantify exposure in an instructed setting, Muñoz (2014) investigated the relationships between starting age, L2 input, and the oral performance of

L2 English language learners from Spain. The researcher found empirical evidence suggesting that the quantity and quality of L2 input (specifically, through the contact with L1 speakers) were stronger predictors of oral performance than starting age. Muñoz also pointed out that "intensive exposure seems to be more effective than long periods of drip-feed instruction" (p. 16), which is consistent with ISLA studies conducted in immersion programs (e.g., Merrill et al., 2021; Pastushenkov & McIntyre, 2020). Methodologically speaking, Muñoz emphasized the importance of breaking down learning experiences into specific facets and measuring exposure by the number of hours/years (e.g., years of instruction, number of hours of curricular and extracurricular activities, social immersion context abroad, informal out-of-school contact with the L2, etc.) of L2 experience.

In terms of the development of L2 speech, L2 researchers have pointed out that some aspects of fluent speech can be learned early in the process of L2 learning, while more nativelike L2 speech characteristics require years of exposure (e.g., Trofimovich & Baker, 2006), emphasizing the need for longitudinal ISLA research in this area. An example of longitudinal L2 speech research that also considered L2 exposure (i.e., the total hours of classes and extra L2 use outside these classes) was the study by Saito and Hanzawa (2018) conducted over the course of one academic year at a Japanese university. The researchers adopted a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018). In addition to recording the first and second semesters to assess the number of hours that the participants spent learning English in their formal classes and informal activities. The researchers then conducted interviews to assess the nature of L2 input that the participants received (form-focused vs. content-based), their level of understanding of their instructors, their engagement in pair work, and other aspects

related to English language pedagogy. The researchers pointed out that the input-proficiency link was particularly strong during the initial phase of the development of L2 fluency and lexicogrammar. Another important finding was that some participants intentionally increased the amount of L2 exposure by enrolling in optional content-based classes. Saito and Hanzawa pointed out the amount of input that the participants received in these optional classes was linked to their L2 speaking gains. Remaining consistent with previous DST literature discussed earlier in this chapter (e.g., de Bot et al., 2007; Ellis, 2007; Hiver, 2015; Larsen-Freeman, 2014), Saito and Hanzawa (2018) concluded that the amount and type of input that their participants received, as well as their pronunciation development, were continuously changing over the course of two semesters, suggesting that a longitudinal perspective is vital in this type of research; this was also realized in their more recent work (e.g., Hanzawa, 2021).

Remaining consistent with DST principles (e.g., Waninge et al., 2014), it is also important to consider the type of the instructed context (Freed et al., 2004; Isabelli-García & Lacorte, 2016; Saito, 2015), as well as individuality of each learner (Ranta & Meckelborg, 2013) in an ISLA study of L2 exposure. For example, Freed et al. (2004) found that the students from an intensive summer immersion program reported that they spent more hours per week speaking and writing L2 French than the students from a regular classroom at a domestic institution and in a study abroad setting: 79.41 hours in the immersion program, 26.36 in the study-abroad program, and 10.85 hours in the traditional classroom. In a study of L2 exposure in different contexts, Martinsen et al. (2011) compared L2 speaking gains of classroom-only students and students living in foreign language housing, which aimed to provide a domestic immersion experience for L2 learners. To assess their participants' L2 exposure, the researchers used a language log, which has become a commonly used instrument in ISLA studies of L2 exposure

(e.g., Pastushenkov, 2020; Ranta & Meckelborg, 2013). In this log, Martinsen et al. (2011) asked the students to record "the number of minutes that they spent speaking, reading, listening to, or writing" in their L2 in the following categories: getting ready for school/work; eating breakfast; in classes; eating lunch; talking to friends/a roommate; watching TV; listening to music; preparing dinner; eating dinner; working; cleaning; studying/doing homework; email; internet; reading; talking on the phone; Sunday school; family home evening; at work; Teaching Resource Center; and "Study Buddy" (p. 280).

Overall, Martinsen et al. (2011) found that the participants used their L2 (French, German, Japanese, and Russian) more in the foreign language housing context than their peers in the classroom-only group, particularly when it came to social situations (e.g., eating meals, preparing meals, etc.), as well as personal time when using the L2 was not necessarily expected or required (e.g., reading email, browsing the internet, listening to music, etc.). In other words, the students felt comfortable using their L2s in different situations, as well as being able to "simulate an immersion experience with each other" (p. 285), which eventually transformed into larger oral proficiency gains compared to the classroom-only group. This finding is consistent with other research regarding the link between the amount of L2 exposure and proficiency gains (e.g., Pastushenkov, 2020; Saito & Hanzawa, 2018). The study by Martinsen et al. (2011) also emphasized the need to consider L2 exposure from both formal classes and additional informal sources, remaining consistent with the propositions from other studies (e.g., De Wilde et al. 2022; Leona et al., 2021; Muñoz, 2014; Saito & Hanzawa, 2018).

With regards to the use of L2 exposure logs, it is important to mention the study by Ranta and Meckelborg (2013). By adopting a cognitivist perspective (while suggesting the importance of sociopsychological and sociocultural dimensions), the researchers found a notable variation in

the amounts and types of language use that their participants (Chinese graduate students at a Canadian university) engaged it, emphasizing the importance of input, output, and interaction as one of the most important predictors of successful SLA. Ranta and Meckelborg also identified a general trend toward receptive rather than interactive use of the L2 and suggested that even if an English language learner lived in an English-speaking country, they did not necessarily actively interact with others in English. Pedagogically speaking, this finding indicated that it may be important to create more opportunities for L2 interaction for learners, specifically outside formal classes. As for future directions, Ranta and Meckelborg (2020) argued that longitudinal studies with multiple data points are needed to effectively assess the effects of exposure, which is also consistent with DST principles (e.g., Serafini, 2017).

IDs in WM, motivation, and L2 exposure as a complex dynamic system

Propositions regarding the interconnectedness of conceptually different categories of IDs (e.g., Dörnyei, 2010; Dörnyei & Ryan, 2015; Larsen-Freeman, 2014; Segalowitz & Trofimovich, 2012) have been realized in ISLA studies at the intersection of cognitive and affective IDs (e.g., Serafini, 2017; Winke, 2013), as well as in research focusing on IDs in cognitive abilities (primarily WM and L2 aptitude) and L2 exposure, specifically the amount of L2 input (e.g., Denhovska et al., 2016, Pattemore & Muñoz, 2020). As Dörnyei (2010) pointed out, "the separation of cognition and motivation has been increasingly seen as an outdated and inaccurate conceptualization in cognitive and educational psychology, let us return to the question of adopting a new, dynamic systems perspective on individual differences" (p. 259). The view of conceptually different ID categories as multi-layered constructs in a complex and dynamic system (e.g., de Bot et al., 2007; Ellis, 2007; Waninge et al., 2014) are comprised of different interconnected elements that need to be studied longitudinally, with multiple data points aligning

with the propositions discussed earlier in this chapter in the sections focusing on WM (e.g., Jackson, 2020, Sagarra, 2017, 2019; Serafini & Sanz, 2016), motivation (e.g., Boo et al., 2015; Gardner et al., 2004; Papi & Hiver, 2020), and additional L2 exposure (Martinsen et al., 2011; Pastushenkov, 2020; Ranta & Meckelborg, 2013).

Despite the rich insight provided by research at the intersection of IDs in cognitive abilities/motivation/L2 exposure (e.g., Denhovska et al., 2016; Pattemore & Muñoz, 2020; Serafini, 2017; Winke, 2013), this area of research remains underexplored in the domain of ISLA. By addressing this gap, empirical ISLA research guided by DST (e.g, Papi & Hiver, 2020; Serafini, 2017) and multidimensional perspectives on L2 learners' IDs (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013) have the potential to inform both L2 theory and pedagogy. A broader holistic approach to IDs (e.g., de Bot et al., 2007; Ellis, 2007) can help reconsider the cognitive/affective dichotomy (e.g., Dörnyei, 2010, Serafini, 2017) and better support L2 learners in different contexts by taking their IDs into consideration when making pedagogical decisions (see, e.g., Csizér, 2017; Li, 2017; Wiley et al., 2014). In one of the empirical studies that simultaneously considered multiple IDs, Winke (2013) used structural equation modeling to investigate the relationships between L2 aptitude, WM, motivation, and strategic use of American military personnel learning L2 Chinese at an advanced proficiency level. The need to investigate IDs and L2 development at an advanced level has also been pointed in other studies of LCTLs, including Russian ISLA research (e.g., Merrill et al., 2021). Winke found evidence that L2 aptitude, which was defined as a multi-layered construct including rote memory, phonetic coding ability, grammatical sensitivity, and phonological WM, was a moderately useful ID construct in this L2 learning context and proficiency level, further emphasizing the need to consider the role of context (e.g., Dörnyei, 2009b; Faretta-Stutenberg & Morgan-Short, 2018;

Larsen-Freeman, 2014; Papi & Hiver, 2020; Pattemore & Muñoz, 2020) and L2 proficiency (e.g., Pastushenkov, 2020; Sagarra, 2017; Serafini, 2017) in ID research, which also aligned with DST principles (e.g., Waninge et al., 2014).

Winke (2013) suggested that the effects of L2 aptitude were mediated by the participants' affective IDs (motivation and strategy use), which further supported the need to reconsider the cognitive/affective dichotomy (e.g., Dörnyei, 2010). Winke did not find strong empirical evidence of a facilitative role for aptitude in explaining attainment variance in advanced learners of L2 Chinese. Phonological WM was the least impactful component of L2 aptitude, with aptitude, strategy use, and motivation having similar positive effects on learning, but different effects with regards to how well they predicted the learners' individual listening, reading, and speaking skills. This finding suggested the need to focus on multiple languages skills in an ISLA study, which has not been a common practice in previous ID research as illustrated in this literature review. The researcher also suggested that the factor loading of WM in L2 aptitude was surprisingly low and that future research would need to administer other WM tests (see "WM in ISLA" section for more detail), in addition to the listening span task used in the study. Moreover, the researcher argued that more work would be required to adapt and refine motivation and strategy use instruments (see "Motivation in ISLA" section for more detail). Pedagogically speaking, Winke (2013) stated that successful L2 development was a combination of various factors and that "the learner needs excellent instruction, frequent opportunities for different kinds of output, and a heavy dose of motivation" (p. 122), further emphasizing the pedagogical potential of this type of research.

The idea of the interconnectedness of learners' IDs (e.g., WM and L2 exposure) has also been discussed in Russian ISLA research. In their laboratory-based studies, Denhovska and

colleagues (Denhovska & Serratrice, 2017; Denhovska et al., 2016) utilized the complexity of Russian morphology and explored the role of WM and L2 input in the incidental learning of noun-adjective agreement by learners having no prior knowledge of Slavic languages. Denhovska et al. (2016) found evidence that at the early stages, the higher WM capacity helped learners produce the target morphological form correctly; however, WM was not engaged when production was supported by high input frequency. For future research, Denhovska and colleagues pointed out the need to test their findings longitudinally in ecologically valid instructed settings. The findings of Denhovska et al. (2016) were consistent with the results of the meta-analysis by Linck et al. (2014), who stated that "the variability in education, L1 abilities, and length of L2 exposure may have introduced additional noise into the data that attenuated any detectable relationship between WM and the outcomes" (p. 878).

In terms of pedagogical implications, ISLA research at the intersection of WM/L2 exposure can help develop strategies of how to better support learners with different cognitive capacities. For example, in their study of learning L2 construction from captioned audio-visual exposure, Pattemore and Muñoz (2020) suggested that "students with a lower WM capacity may need the support of captions as well as other types of support (e.g., focus on form) to benefit from exposure to multimedia" (p. 9), remaining consistent with the propositions of other researchers (e.g., Wiley et al., 2014). Indrarathne and Kormos (2018) also found empirical evidence that L2 learners with a higher WM capacity improved their receptive knowledge more successfully than L2 learners with a lower WM capacity. According to the researchers, this finding was important for L2 pedagogy, as it suggested that L2 learners with a lower WM capacity may require extensive L2 exposure to grammar, as well as additional instructional support. With regards to the benefits of additional L2 exposure, Leona et al. (2021) pointed out

in their study at the intersection of L2 exposure/motivation that the role of motivational frameworks (socio-education model of SLA or L2MSS) did not help explain L2 performance in young English language learners, whereas the sources of extramural English exposure helped predict the learners' performance on both oral and written receptive vocabulary tests. In line with other studies discussed in this chapter (e.g., De Wilde et al. 2022; Martinsen et al., 2011; Muñoz, 2014; Saito & Hanzawa, 2018), Leona et al. (2021) emphasized the need to distinguish formal and informal L2 exposure. Despite the rich insight provided by their work, it is important to note that the participants in the study by Leona et al. (2021) were children and thus these findings were not necessarily generalizable to adult L2 learning.

Another ISLA study that simultaneously investigated conceptually different ID categories (L2 exposure, memory, age of onset, and motivation) in different contexts (foreign language in Poland and immersion in the UK) was the work done by Foster et al. (2014). The researchers of this study investigated the role of these IDs and contextual variables in the receptive knowledge of native-like selections (NLS) of English language learners. Using hierarchical regression analysis, Foster et al. (2014) found empirical evidence that the learners' age of onset was the strongest predictor of native-likeness. The researchers also found that higher L2 motivation was not associated with native-likeness and that phonological short-term memory was the only predictor of NLS ability in immersion late starters, further emphasizing the need to consider the role of context in a study of IDs (see, e.g., Martinsen et al., 2011). With regards to the length of L2 exposure, Foster et al. (2014) pointed out that "longer exposure brings moderate improvement to NLS for [non-native speakers] NNSs living within the [target language] TL community but is not sufficient to bring participants to nativelike levels without the benefit of an early start" (p. 121). Remaining consistent with the multidimensional perspectives of IDs (e.g.,

Sun & Zhang, 2020; Wang & Eccles), the study by Foster et al. (2014) further illustrated the need to consider multiple IDs in a study of L2 development in instructed settings.

One of the few longitudinal studies that attempted to apply a dynamic system lens and viewed conceptually different IDs as interconnected constructs in instructed L2 learning was the work by Serafini (2017), which was one of the studies that guided this dissertation's theoretical orientation and methodology and thus is described in more detail. The participants (N = 87) were enrolled in beginning, intermediate, and advanced L2 Spanish courses at a university in the United States. The students were enrolled in a non-intensive Spanish program and received 150 minutes of instruction per week (50 minutes per day, three days per week) for approximately 14 weeks per semester. Serafini pointed out the importance of teacher variables, as the participants were recruited from different classes and classroom interaction was not observed. The instructors, however, were educated from a psycholinguistic perspective and were encouraged to use both explicit and implicit approaches to L2 pedagogy. Serafini also collected additional information about the teachers' backgrounds and their practices. The results indicated that the teachers used similar activities (e.g., listening, speaking, reading, and writing tasks; technologymediated tasks; individual/pair/group work; and teacher-led discussions), balanced the use of explicit and implicit feedback, and strove to promote a student-centered classroom.

Serafini adopted two measures of WM capacity: an OSpan measuring the executive function component of L2 learners' WM capacity, and a Digit Span task measuring their phonological WM. In line with previous research (e.g., van den Noort et al., 2006), the WM tasks⁴ were conducted in the participants' L1 (English). Both tasks were reported to have high reliability in research with similar college-age participant populations (e.g., Conway et al., 2005).

⁴ Similar WM tasks were used in the present study and are discussed in detail in the next sections of this dissertation.

In addition to the two WM tasks, Serafini (2017) adopted two measures of L2 motivation: the AMTB developed by Gardner and colleagues (e.g., Gardner & Lambert, 1972, Gardner et al., 2004; Masgoret & Gardner, 2003), and Dörnyei's L2 Selves questionnaire (e.g., Dörnyei, 2009b; Dörnyei. 2010). The participants self-reported their motivation using a 7-point Likert scale. The AMTB included 78 items and the following sub-scales: integrativeness (i.e., integrative orientation, interest in foreign language learning, and attitudes toward the L2 community); attitudes toward the learning situation (i.e., attitudes toward the teacher and the course); and motivation (i.e., effort, desire, and attitude toward learning). The L2 selves motivation questionnaire included 30 items on a 7-point Likert scale. The participants self-reported their motivation assessed using this questionnaire in the following categories: the ideal L2 self (i.e., characteristics that an L2 learner would like to have); the ought-to L2 self (i.e., characteristics that L2 learners think they ought to have); and motivated learning behavior (i.e., the effort learners intend to put in L2 learning).

The two WM tasks were completed once at the beginning of the semester; the motivation surveys were distributed twice at the beginning and at the end of the semester (the time difference between the two phases of the study was 2.5 months). Serafini pointed out that the design with only two data points was considered a limitation, as it did not necessarily differentiate the study from traditional ISLA research. In the analyses, the researcher calculated descriptive statistics for ID measures and checked reliability of the motivation instruments using Cronbach's alpha coefficients. To address the study's research questions (RQs), Serafini conducted bivariate correlation analyses and interpreted the results based on the statistical significance and the Pearson's r effect size: small (r = .10), medium (r = .30), and large (r = .50). The data were also explored visually using simple scatterplot and scatterplot matrices with linear

and Loess lines (this helped highlight linearity or non-linearity of the data). If relationships between IDs changed over time (from conditional to supportive), these relationships were considered asymmetric.

The results of the bivariate correlation analyses showed several instances of meaningful relationships. In the advanced group at the outset of instruction (Time 1 or Phase 1), two negative relationships were found between WM as measured by OSpan, as well as motivation intensity ($r = -.37^*$, Power⁵ = 53%) and the ought-to L2 self ($r = -.39^*$, Power = 58%), suggesting that a lower ability to store and process information simultaneously as measured by OSpan was associated with more effort (i.e., motivation intensity) and a stronger ought-to L2 self. At the intermediate level, the L2 Spanish learners' ability to code and store verbal information assessed by the phonological Digit Span task was associated with a stronger identification (i.e., integrativeness) with the Spanish-speaking community ($r = .38^*$, Power = 59%) and a more positive ideal L2 self ($r = .39^*$, Power = 62%). A statistically significant correlation was found at the advanced level between the composite scores on the OSpan task (referred to as executive function) and language anxiety, which was another component of AMTB ($r = .46^*$, Power = 73%). At the beginning proficiency level, a statistically significant correlation was found between the phonological component of WM and the participants' ideal L2 self ($r = .45^*$, Power = 53%). At Time 2 (or Phase 2), 2.5 months after Time 1, four more significant correlations were found. Serafini pointed out that the negative relationship between executive function and integrativeness for advanced learners endured from Time 1 (r = -.44*, Power = 69%) to Time 2 (r = -.38*, Power = 53%), meaning that lower scores on the OSpan

⁵ "A power level of .80 indicates a reliable effect in the population whereas a power level of .50 indicates a 50% risk of failing to confirm a valid research hypothesis" (Serafini, 2017, p. 378).

task were consistently related to a stronger desire of L2 learners to identify with the Spanishspeaking community. Another association that had maintained over the course of the study was observed in the correlations between phonological WM and language anxiety for beginner L2 learners of Spanish (from r = .44*, Power = 52% to r = .46*, Power = 55%). For the intermediate group, higher scores on the phonological Digit Span task were associated with a stronger effort and/or desire to learn Spanish as a L2 (from r = .48**, Power = 83% to r = .45**, Power = 74%), as well as with more favorable attitudes toward the teacher and/or the course (from r = .44*, Power = 72% to r = .40*, Power = 61%).

Serafini also pointed that the relationships between elements in two subsystems (interpreted as 'growers') were supportive, conditional, competitive, and compensatory, and such relationships changed over time and varied at different proficiency levels. For example, the researcher found evidence that a stronger integrative motivation and/or more effort (i.e., motivational intensity), along with a more positive ought-to L2 self compensated for a lower WM capacity as measured by the OSpan task in the advanced group. Serafini also discussed the emergence of so-called attractor states, or points of stability. For example, at the intermediate level, all relationships between the IDs were considered supportive, while most of the relationships ranged from conditional to supportive. In conclusion, Serafini (2017) started that this study has illustrated that "DST provides a larger, more holistic framework necessary for interpreting complex, shifting interrelationships of learner resources in an instructed setting and offers the conceptual tools needed to broaden and deepen our understanding of the cognition–motivation interface in adult L2 development" (p. 384).

IDs in Russian ISLA: Future directions

Despite the growing interest in LCTLs (e.g., Kim, 2017; Pastushenkov & McIntyre, 2020; Thompson 2017; Winke, 2013), including Russian (e.g., Cho & Slabakova, 2014; DeKeyser, 2012; Klimanova & Dembovskaya, 2013; Merrill, 2020; Sudina & Plonsky, 2021), this area remains relatively understudied within the domain of ISLA research.⁶ As it was illustrated in the earlier sections of this literature review, most of the ISLA studies of IDs investigated the L2 learning of more commonly taught languages such as English (e.g., Foster et al., 2014; Leona et al., 2021; Papi & Hiver, 2020) and Spanish (e.g., Grey et al., 2015; Sagarra, 2017; Serafini, 2017). Previous studies in the area of Russian ISLA have often focused on different aspects of the complex Russian grammar, particularly its rich inflectional morphology (see, e.g., Kisselev et al., forthcoming). Various researchers have studied the L2 acquisition of Russian case markings (e.g., P. Brooks & Kempe, 2013) and gender agreement (Denhovska & Serratrice, 2017; Denhovska et al., 2016). Even fewer studies have looked at the experiences of Russian learners, including why they chose to study the language (e.g., Merrill, 2013, 2020), their motivation dynamics (e.g., Pastushenkov & McIntyre, 2020), their beliefs about study abroad (e.g., Zaykovskaya et al., 2017), or their use of Russian social media (e.g., Klimanova & Dembovskaya, 2013). Considering the geopolitical situation in the 1980s and the early 1990s and immigration from the countries of the former Soviet Union, heritage language education has been another important topic that L2 Russian scholars have explored (e.g., Kisselev et al., 2020), with heritage language status being considered an important ID in previous ISLA literature (e.g., Montrul & Bowles, 2017).

⁶ It is important to note that previous researchers have explored L2 Russian learning in addition to other languages (e.g., Martinsen et al., 2011; Sun & Zhang, 2020).

With a few exceptions that included relatively large samples of participants, specifically for research with LCTLs (e.g., Kisselev et al., forthcoming), previous studies with L2 Russian learners have often adopted case study methodologies with only one or two focal participants (e.g., Pastushenkov & McIntyre, 2020; Zaykovskaya et al., 2017). Despite the benefits of this learner-centered approach, future Russian ISLA should consider expanding its sample size. As Russian is an LCTL and finding participants can be challenging, a possible solution is to adopt a mixed-methods methodology (e.g., a sequential explanatory design, see Ivankova et al., 2006; Taguchi, 2018), in which qualitative data from interviews or artifacts can help clarify some of the patterns in the quantitative data from questionnaires or other sources (e.g., Pastushenkov & McIntyre, 2020; Saito & Hanzawa, 2018).

Despite the rich insight provided by researchers within the domain of IDs in Russian ISLA, including the studies discussed in the earlier sections of this literature review (e.g., Denhovska & Serratrice, 2017; Denhovska et al., 2016), future research will need to test their findings in more ecologically valid settings, over an extended period of time, and at a wide range of initial proficiency levels. Remaining consistent with the principles of the complexity/DST framework (e.g., Larsen-Freeman, 2014; Papi & Hiver, 2020), the need for longitudinal studies with multiple points has been pointed out by L2 researchers working in areas of WM (e.g., Sagarra, 2019), motivation (e.g., Pastushenkov & McIntyre, 2020), L2 exposure (e.g., Ranta & Meckelborg, 2013), and also in research at the intersection of these ID categories (e.g., Serafini, 2017). Considering that morphologically rich languages such as Russian require extended learning periods for most L1 speakers of English (Looney & Lusin, 2019), a longitudinal perspective is vital for an ID study in the domain of Russian ISLA.

Promoting ecological validity can help address the criticisms of the field of SLA and its relevance and applicability to L2 pedagogy (see Marsden & Kasprowicz, 2017; Sato & Loewen, 2019b; Spada, 2015). As Godfroid (2016) pointed out, conducting research in ecologically valid settings with natural languages and real L2 learners has its challenges (e.g., taking participants' prior knowledge into consideration); however, "... learning does not happen in a vacuum" (p. 207). Remaining consistent with these propositions, various L2 researchers pointed out the need to consider learners' backgrounds and initial L2 proficiencies in studies of IDs (e.g., Sagarra, 2017; Serafini, 2017). To the best of my knowledge, L2 Russian proficiency has only been considered in very few studies (e.g., Cho & Slabakova, 2014). Will the findings of previous Russian ISLA research hold true at different proficiency levels in the long term in an ecologically valid instructed setting? Moreover, to what extent can the findings of previous research be transferrable to the acquisition of more complex lexicogrammatical patterns such as the derivational morphology of Russian verbs of motion (see Pavlenko, 2010) and language skills other than grammar (Pastushenkov, 2020)?

Present study

In line with Serafini (2017), this study aims to further "improve upon conceptual and methodological gaps limiting mainstream learner ID research" (p. 375). Guided by DST (e.g., de Bot et al., 2007; Ellis, 2007; Larsen-Freeman, 2014; Papi & Hiver, 2020) and multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020), this longitudinal mixed-methods study conducted with students from a Russian summer immersion program at four curricular levels aims to explore how L2 learners' motivation and the amount of L2 exposure change over time, as well as how WM, motivation, and the amount of L2 exposure affect learning gains in lexicogrammar, speaking, and writing. This study's primary goal is to promote

research at the intersection of IDs in cognitive abilities, motivation, and L2 exposure that can help obtain a more fine-grained view of how IDs evolve over time and better understand their roles in L2 development. Moreover, by conducting this study with L2 Russian learners from a well-established summer program, I strove to promote the need for ecologically valid ISLA research of LCTLs. This dissertation research is guided by the following research questions (RQs):

RQ 1A: To what extent does L2 Russian students' motivation change at different curricular levels over the course of a summer immersion program?

RQ 1B: To what extent does the amount of L2 Russian exposure from different sources change at different curricular levels over the course of a summer immersion program?

RQ 2: To what extent do L2 Russian students' WM capacity, motivation, and L2 exposure affect their learning gains for lexicogrammar, speaking, and writing at different curricular levels over the course of the summer immersion program?

CHAPTER 3: METHODOLOGY

Russian summer program: Curricular levels and distinct features

The data for this dissertation research were collected at an eight-week Russian summer immersion program in the United States in 2021. Due to the COVID-19 pandemic,⁷ the 2021 program was conducted entirely online and included four curricular levels, in which students were initially placed based on a series of placement tests for lexicogrammar, speaking, and writing discussed later in this chapter. Apart from the summers of 2020 and 2021, the Russian summer program was conducted in-person, including the summer of 2019 when the data for the dissertation pilot study were collected (for more information about the dissertation and test equating pilot studies, see Chapter 4). The summer program ran from the end of June to mid-August.

Before the COVID-19 pandemic, the in-person Russian program had typically included seven curricular levels: Level 1 Introductory Russian, Level 2 Advanced Introductory Russian, Level 3 Basic Intermediate Russian, Level 4 Enhanced Intermediate Russian, Level 5 Advanced Intermediate Russian, Level 6 Advanced Russian I, and Level 7 Advanced Russian II. The data for the dissertation pilot study (Pastushenkov, 2020) were collected in the summer program of 2019, which also included seven curricular levels. The number of curricular levels had important implications for the methodology of the dissertation pilot study conducted in the summer of 2019 and later for dissertation research conducted in the summer of 2021. The formal/informal distinction for Russian L2 exposure (the number of hours that students were engaged in learning Russian) in the summer program are also important for this dissertation research, as one of the

⁷ The dissertation pilot study (Pastushenkov, 2020) was conducted in the face-to-face summer program of 2019. Considering that the dissertation data were collected in the online 2021 program, several modifications in the study's design were made (for more detail, see Chapter 4).

ID categories under investigation is additional L2 exposure (i.e., the amount of L2 Russian learning outside formal classes).

It is also important to note that the program's expectations with regards to the proficiency level that their students reach by the end of the summer program are based on the American Council on the Teaching of Foreign Languages (ACTFL) scale (Novice to Superior), which will be discussed later in this chapter. These estimates are based on the placement and exit testing data from the ACTFL Oral Proficiency Interviews (OPIs)/Oral Proficiency Interviews – Computer (OPIcs) and ACTFL Writing Proficiency Tests (WPTs) that complement the program's internal placement and exit lexicogrammar quizzes. The placement and exit tests conducted by the program and the learning gains associated with these tests are crucial for this study's methodology and are described in detail later in this chapter. Finally, the fact that students often achieve different learning outcomes as seen in the program' estimates for their students' gains by the end of the program (e.g., Novice High to Intermediate Low Russian proficiency at Level 1) is one of the primary motivations for this dissertation research (for more information about the learning gains of L2 Russian students, see Kisselev et al., forthcoming, Merrill et al., 2021 and Chapter 7 of this dissertation).

According to the program's website (the link is not included to maintain the program's anonymity), students at Level 1 are expected to have zero or minimum prior Russian classroom instruction. By the end of the summer program, students at Level 1 are expected to reach Novice High to Intermediate Low Proficiency levels on the ACTFL scale. At Level 2, students have approximately 100 hours of previous formal instruction in Russian (one to two semesters of college Russian at three hours per week). Upon completion of Level 2, L2 Russian learners typically reach Intermediate Low to Intermediate Mid Russian language proficiency.

At initial Curricular Level 3, students in the Russian summer program have

approximately 150 hours of prior formal Russian instruction. One of the primary goals for Level 3 is to acquire an active vocabulary of 1,500 words, which also informed the development of the program's lexicogrammar placement and exit tests. Upon completion of Level 3, students typically reach Intermediate Mid language skills. At Level 4, the students have approximately 200 hours of previous Russian instruction; the goal is to acquire 1,700 words; and students tend to reach Intermediate Mid to Intermediate High proficiency levels.

At Level 5, students have approximately 300 hours of prior formal Russian instruction. The goal for this level is to approach 2,000 Russian words and to be able to read literature in the Russian language and newspaper articles. A significant portion of the summer program at Level 5 is devoted to news. Students typically reach Intermediate Mid to Intermediate High proficiency upon completion of the summer program. At Level 6, students have approximately 350 hours of prior formal Russian classes (or fewer hours but a semester or more that they studied Russian abroad in a Russian-speaking country). At this level, students study more advanced grammatical structures such as participles, verbal adverbs, quantitative expressions, and verbs of motion. They are actively engaged in watching movies in Russian and reading Russian literature as part of their formal classes. In terms of assessment, students complete challenging assignments including journalistic work and oral presentations. Upon completion of Level 6, students usually reach Intermediate High to Advanced Low Russian proficiency.

Finally, at Level 7, students are expected to have approximately 400 hours of prior formal instruction in Russian (or fewer hours if they completed at least one semester of study abroad in a Russian-speaking country). At this level, instructors in the summer program focus on complicated grammatical structures and help students increase their vocabulary by adding

idioms, synonyms, and root-based semantic groups. Developing lexical competence (i.e., an active vocabulary of 2,500 words), as well as mastering prototypical models of word formation and derivation process are primary goals at Level 7. At this level, the program puts an emphasis on engaging in and analyzing contemporary Russian culture (particularly, literature and film). During the program, students at Level 7 give a presentation about Russian-speaking authors. Students at this level usually reach Advanced Low to Advanced Mid Russian language proficiency by the end of the summer program.

In addition to having a wide range of initial curricular levels, another important feature of the Russian program is that it usually recruits a large number of L2 Russian learners. The large number of students makes this summer program stand out compared to traditional Russian classrooms at the majority of universities in the United States and the other immersion programs such as STARTALK that often have small cohorts and focus on beginner learners (see Pastushenkov & McIntyre, 2020). During the summer of 2019 when the data for the dissertation pilot study (Pastushenkov, 2020) were collected, 137 students were enrolled in the program (25 students in Level 1; 14 in Level 2; 21 in Level 3; 22 in Level 4; 26 in Level 5; 14 in Level 6; and 15 in Level 7).

In 2020, the program had to switch to an emergency remote instruction model due to the COVID-19 pandemic and the lower number of students (83 students total). Considering the circumstances, I decided to postpone my data collection to the summer 2021 program and modify the materials and procedure if necessary (for more detail about the modifications, see Chapter 4). Even though the summer program of 2021 was also conducted entirely online, the instruction model was not considered emergency remote, and the enrollment was higher. However, the program only included four curricular levels. According to the official website (the

link is not included to maintain the program's anonymity), the four levels were the following: First-Year Russian, Second-Year Russian, Third-Year Russian, and Fourth-Year Russian (in this dissertation, referred to as Curricular Levels 1, 2, 3, and 4, respectively). From 83 students in the summer of 2020, the number went to 111 students at the four curricular levels in the summer of 2021 (27 students in Level 1; 30 in Level 2; 30 in Level 3; and 24 in Level 4).

As mentioned on the program's website, the placement criteria were somewhat broader for the online program compared to the face-to-face programs that had typically included seven curricular levels. Therefore, in order to accurately assess the relationships between IDs and the learning gains at individual curricular levels, I verified that the students enrolled in these levels had similar scores on the placement tests (for more detail, see Chapter 7). Curricular Level 1 was designed for students with zero or minimum prior Russian classroom learning instruction⁸. At this level, the students learned the Russian alphabet and how to read and write in Russian. The students also learned how to speak and understand spoken Russian in basic and predictable realworld contexts such as ordering food at a restaurant or asking directions. At Level 1, Russian language instruction focused on basic grammatical structures and the acquisition of beginning vocabulary, which the students also practiced in weekly compositions. Upon completion of this level, the students were expected to move to Novice High or Intermediate Low Russian proficiency on the ACTFL scale (for more information about the students' progress in the summer program, see Chapter 8).

Curricular Level 2 was designed for students who had approximately 150 hours of prior classroom instruction in the Russian language. At this level, the students reviewed basic

⁸ The fact that some of the students had zero experience learning Russian was taken into consideration in the design of the L2MSS survey that included Russian learning experience in addition to the three types of selves (the materials are discussed in detail later in this chapter).

grammatical structures to have a strong foundation for their future studies, as well as expanded their active vocabulary to 1,500 words. Upon completion of the summer program, the students were expected to reach Intermediate Mid proficiency in Russian. At Curricular Level 3, the students who were initially placed in this level had approximately 300 hours of formal Russian learning in instructed settings. The instruction at this level focused on more difficult grammatical structures of the Russian language such as verbs of motion, verbal aspect, and complex syntax. The goal was to have 2,000 words in the students' active vocabulary. According to the program's website, one of the other important goals was to be able to discuss important social and cultural events and issues in Russian-speaking countries. At this level, the faculty provided in-depth support regarding how these events and issues were covered in contemporary media, literature, and film. Upon completion of Level 3, the students were expected to reach Intermediate Mid to Intermediate High language proficiency.

Finally, at Curricular Level 4, the students had approximately 400 hours of previous Russian learning experience in instructed settings (fewer hours were accepted when the student completed a semester or more of study abroad in a Russian-speaking country). At this level, the instruction focuses on more complicated grammar and aims to increase the students' active vocabulary by adding idioms, synonyms, and root-based semantic forms. In general, the development of lexical competence and being able to use 2,500 words is a major goal at this level, in addition to being familiar with prototypical models of word formation and derivation process in Russian. The students at Curricular Level 4 analyze Russian films and read contemporary Russian literature. Upon completion of Level 4, the students were expected to reach Advance Low to Advanced Mid Russian language proficiency.

Over the course of its history, most of the L2 Russian learners in the summer program were undergraduate and graduate students from the United States; however, students from Europe and other regions often joined the program. From the pilot study (Pastushenkov, 2020) and my visit to the program in the summer of 2019, I learned that some of the students completed the program more than once at different initial curricular levels. It is important to note that the institution also offered a separate six-week program for graduate students only and a four-week Refresher Course during the summer of 2021. For the purposes of this study, data were collected from participants of the eight-week immersion program who were required to take the program's placement and/or exit tests (the program does not conduct entrance placement tests to absolute beginners).

After one of my consultations with the program director, I learned that the placement in the eight-week immersion program depends solely on language proficiency level and not the students' status outside of the summer program (e.g., undergraduate or graduate level). The placement of students in the six-week program also depends on their language proficiency. The program accepts students with a grammar score of 60 (/100) on the lexicogrammar fill-in-theblank quiz and at least Intermediate High on an oral interview. The program looks carefully at those who score Intermediate High, and some graduate students (especially if they received a low score on the writing test) are referred to the eight-week program. The program started the four-week online refresher only in response to COVID-19. Before the pandemic, the refresher was a one-week on-campus program. According to the program director, this practice was very successful and in 2022 the program follows the 2021 model with the section being full as of early June 2022.

Another distinct (and probably the most well-known) feature of this summer program is the language pledge to speak only Russian that all students take prior to starting their studies. Considering how challenging this pledge can be (particularly at lower proficiency levels), to be prepared to speak only Russian during the summer program, students at Level 1 usually start learning the Russian alphabet and survival phrases online one month before the program begins. A modified version of the language pledge was used in the 2021 online summer program, in which the students promised to use as much Russian as possible.

It is important to mention that my frequent consultations with the program director and faculty before and during my visit to the program in 2019 and over the course of online data collection in 2021 and dissertation write-up in 2021-2022 helped me with my dissertation research and provided unique insights into the program. The conversations with the students during the summer of 2019 and my interviews in the Fall of 2021 also helped me better understand the learners' daily routines and prepare the materials that I used in my dissertation. For example, I learned that the vast majority of the students in the summer of 2019, one of the language pledge very seriously. When I visited the program in the summer of 2019, one of the students mentioned that they asked their parents to translate their text messages from English into Russian using Google Translate. I have also noticed some minor violations of the language pledge, probably at lower proficiency levels (for more detail about why students sometimes switched to their L1s and violated the language pledge, see Merrill, 2020).

As mentioned earlier in this chapter, the summer program puts an emphasis on both language and culture. This distinct feature of the program can be observed in both formal classes and students' co-curricular activities. In line with the pre-pandemic tradition that was maintained to a certain extent during the online programs of 2020 and 2021, the students in the summer

program of 2022 can participate in the following co-curricular activities: theater, choir, soccer, volleyball, healthy spine club, daily newspaper / journalism club, political news club, cooking club, traditional painting club, Russian and Soviet rock club, chess club, Russian fair, quiz club, open mics, Russian etiquette, and Russian slang. The first face-to-face summer program after the pandemic also has its own YouTube channel, which includes videos of clubs and guest lectures, as well the productions of the theater and choir.

Russian summer program: Schedule and activities

Based on the program's description on the official website, each curricular level in the eight-week summer program was valued at 9 credit hours and participants met for at least 15 synchronous hours per week (approximately 120 hours of instruction from the end of June to mid-August). This number is important because L2 Russian exposure (specifically exposure from outside formal classes) is one of the three ID categories under investigation. Based on my conversations with the program's faculty and staff, the Russian instructors were better prepared for online teaching in the summer of 2021 after the transition to emergency remote instruction in the summer of 2020 and after teaching online during the Fall 2020 and Spring 2021 semesters at their home institutions.

The instructors' goal was to prepare engaging and enriching activities, to which I will return in Chapter 8 in the discussion of the students' motivation during the program. The program included a combination of synchronous instruction, conversation practice, and cocurricular activities that aimed to improve the students' language proficiency as fast as possible. It is important to note that some of the co-curricular activities (a major component of the face-toface Russian program before the pandemic) were more difficult or even impossible to implement online. For example, before the pandemic, the students in the summer program could play sports

on campus while speaking Russian or participate in the Russian theater club. Despite these limitations, the program strove to maintain its teaching philosophy, including the focus on work in small groups and giving students individualized attention and opportunities for conversations. Such conversations gave students opportunities to practice new vocabulary and grammar.

To help give students individualized attention, the program hired a group of tutors who helped students outside of their formal classes to even out their language skills in speaking, listening, reading, and writing; to fill in the gaps in the knowledge of Russian if necessary; and to practice their phonetic skills. In line with the previous summer programs, the curriculum in the online program of 2021 also covered approximately one academic year of Russian. Outside of their formal classes, the students had three types of Russian practice organized by the program: office hours and tutoring; cultural events that included guest lectures, journalism workshops, and thematic clubs on different cultural topics such as Russian cartoons and internet culture; and informal socialization opportunities using the online tools Zoom and Discord, which were also used for the formal classes (in addition to Canvas, the program's learning management system).

Participants

The participants in the present study (N = 52 out of 111 students in the summer program) were recruited via email with the help of the program director and assistant director using the internal emailing system. The participants were asked to provide their names and email addresses⁹ in Qualtrics, so I would then be able to connect the data from different phases, send the participants reminders about the new phases of the study, and later recruit participants for post-program interviews conducted during the Fall 2021 semester. The numbers of participants

⁹ Each participant was given an identification code to protect their confidentiality; the data files are stored with the codes; and the participants' names and email addresses were deleted upon completion of the project.

in different curricular levels and phases of the study are shown in Table 1. The students were recruited from all four levels. Considering the very low number of participants in Curricular Level 4, particularly on Phases 3, 4, and 5, this level was excluded from some parts of the analyses (see Chapters 6 and 7).

Table 1. Number of participants at different curricular levels and phases of the study					
	Curricular	Curricular	Curricular	Curricular	TOTAL
	Level 1	Level 2	Level 3	Level 4	
Phase 1	14	14	15	9	52
Phase 2	10	10	10	4	34
Phase 3	9	11	10	3	33
Phase 4	8	11	7	3	29
Phase 5	7	8	7	2	24
Post-program interviews	4	3	2	2	11

Table 1. Number of participants at different curricular levels and phases of the study

On average, the participants were 24.12 years old (SD = 7.09) and started learning Russian at the age of 21.10 years old (SD = 5.18). The students identified as females (n = 32), males (n = 20), other (n = 1), and one participant who preferred not to disclose their gender. The sample included undergraduate (n = 25) and graduate students (n = 21), as well as L2 Russian learners who were self-reported "summer students" (n = 6). Most participants in the program were L1 speakers of English (n = 42). The sample also included one L1 speaker of Italian, one L1 speaker of Portuguese, and two L1 speakers of Spanish. One participant reported to be a bilingual (L1 speaker of English and French); another participant was self-reported trilingual (L1 speaker of Mandarin, English, and Taiwanese). There was also one self-reported L1 speaker of Russian.¹⁰ Upon my consultation with the program director and faculty and after conducting the post-program interviews in English, I verified that the non-L1 speakers of English were at

¹⁰ From the post-program interview with this participant, I learned that this student was born in Russia but moved to the United States a child and no longer spoke Russian with their adoptive family.
advanced and/or near-native/native proficiency levels in English (which was important for the WM tests used in the study, for example). I also analyzed the data for potential outliers and went through several steps of data cleaning (for more information about these procedures, see Chapter 5). Considering that the sample size was relatively low (particularly at individual curricular levels and later stages of the study, see Table 1) and that the sample was generally representative of the population of students in the Russian summer program, I decided not to exclude any participants from the analyses based on their L1 background.

The students were compensated for their participation in the study by using funding from the *Language Learning* Dissertation Grant (The Language Learning Dissertation Grant Program, USA) and the NFMLTA / NCOLCTL Dissertation Grant (The National Federation of Modern Language Teachers Associations and The National Council of Less Commonly Taught Languages, USA). The participants were given gift cards of a major online retailer (up to \$40): \$10 after completing Phase 1; \$20 after completing Phases 2-5; and \$10 for the interview. Upon completion of Phases 1-5, the participants were also given 60 minutes of free Russian tutoring during the Fall 2021 semester.

Placement and exit tests

The summer program's lexicogrammar, speaking, and writing placement and exit tests were used to calculate learning gains (for more information about the program and its tests, see Merrill et al., 2021). The program considered the tests to be valid instruments for placement purposes. It is also important to note that SLA researchers have previously found misfitting items in the ACTFL Can-Do Statements that align with the Novice-Superior proficiency scale and pointed out that some of the statements tended to be vague (see Tigchelaar et al., 2017). The placement and exit tests have similar structures to ensure that they are equally difficult and can

60

be used to reliably measure learning gains for lexicogrammar, speaking, and writing (for more information about test equating pilot study, see Pastushenkov, 2020 and Chapter 4 of this dissertation). The tests included two online lexicogrammar quizzes (fill-in-the-blank [FIB] and multiple-choice [MC]), speaking tests (the ACTFL OPIcs), and writing tests (the ACTFL WPTs). The tests are not publicly available. The entrance tests were not distributed to absolute beginners.

The OPIcs were conducted online before and at the end of the summer program and rated by certified ACTFL examiners who assigned ratings on the ACTFL scale (see Figure 3). The tests take 20-40 minutes to complete and cover topics related to everyday life, school, and work, as well as current events. According to the official website of ACTFL, the online OPIcs emulate the 'live' OPI (a 15-30 telephone conversation), which had been used in the summer program before 2021. One of the key differences between the 'live' OPIs and OPIcs is that the latter delivers prompts through a computer program and uses a virtual avatar named Ava. The OPIcs use the following format: introduction and warm-up; background survey; self-assessment; and forms. During the introduction and warm-up, test takers are given an explanation of the procedures and instructions of the OPIcs and a sample speaking prompt (e.g., tell Ava about yourself). The background survey asks L2 learners questions about their educational and professional background, as well as about their favorite activities and interests. This survey helps the program determine the pool of prompts that will be used in the OPIc. Upon completion of the background survey, the participants are given a self-assessment in which test takers describe their speaking ability. This self-assessment determines which of the five versions (or forms) of OPIcs will be given to the L2 learner: Form 1 for the Novice Low to Intermediate Low range; Form 2 for the Novice Low to Intermediate High range; Form 3 for the Novice Low to Advanced Low range; Form 4 for the Intermediate High to Advanced High range; and Form 5 for the Advanced Mid to Superior range. Some of the assessment factors include the L2 learners' ability to complete the speaking task, how well they can support their opinions, how well the raters understood the test taker, their pronunciation, etc.



Figure 3. ACTFL inverted pyramid representing the functions of the major levels (actfl.org).

Similarly to the OPICs, the WPTs were also rated using the ACTFL scale (see Figure 3). The writing tests were conducted before and at the end of the summer program (for information about the WPTs used in the program and how the ratings were assigned, see Kisselev et al., forthcoming; Merrill et al, 2021). Since the program was conducted entirely online, the students' writing tests were not supervised by proctors, which is considered a potential limitation even though the students were not given grades and the tests were used for placement purposes and to track the students' progress over the course of the summer program. The students were given up to four writing prompts and 60 minutes for each WPT. Some of the topics that the students wrote about were politics, Russian learning experience, global challenges, etc. Upon completion of the OPIcs and the WPTs, the students in the summer program were assigned one of the proficiency levels on the ACTFL scale (for the functions of each major level, see Figure 3). The scale includes four main levels (Novice, Intermediate, Advanced, and Superior) and three sublevels (Low, Mid, and High). Based on my previous experience and research conducted in the summer program (e.g., Kisselev et al., forthcoming; Merrill et al, 2021), students usually joined the program at Novice, Intermediate, and Advanced Levels, but there were also cases when L2 Russian learners reached Superior proficiency¹¹ by the end of the program.

For the purposes of this study (specifically RQ 2 described in Chapter 7), the ACTFL scale (Novice to Superior) was converted into a numeric 0-9 scale (see, e.g., Kisselev et al., forthcoming; Merrill et al., 2021; no participants in this study reached the superior level on the ACTFL scale and thus the 0-10 scale was not used). As the absolute beginners did not complete the placement tests, they were assigned 0 on the numeric scale in line with Kisselev et al. (forthcoming). A score of 1 corresponded to Novice Low; 2 to Novice Mid; 3 to Novice High; 4 to Intermediate Low; 5 to Intermediate Mid; 6 to Intermediate High; 7 to Advanced Low; 8 to Advanced Mid; 9 to Advanced High; and 10 to Superior. After gaining access to the testing results for 2021, I noticed that some of the scores had a plus sign after them (e.g., Novice Low+).

¹¹ ACTFL Superior does not have sublevels.

I consulted with one of the experienced Russian instructors and raters from the program. I learned that for the program's internal purposes, the raters added a plus to the ACTFL rating if the participants were in-between the two levels, which was often the case for intermediate and advanced students. I decided to take this procedure into account when transferring the OPIc and WPT scores into the numeric scale. If the raters in the summer program added a plus to their ratings, I added 0.5 to the numeric scale (e.g., Intermediate Low+ would correspond to 4.5), as it would provide a more fine-grained view of the participants' L2 development in the summer.

The program had delivered their lexicogrammar quizzes online even before the COVID-19 pandemic, so the 2021 program did not have to modify the procedure for their internal tests. Students can complete the quizzes on their own electronic devices. Each test has a time limit of 60 minutes and covers a wide range of grammatical and lexical forms. The placement and exit versions of the tests are different; however, students at all levels are given the same tests for placement and the same tests for exit (the tests are non-adaptive). Before the students start the tests, they are informed that they may not be able to answer all the questions. There are no grades; the tests are conducted for placement purposes and to measure the students' progress in the summer program. Based on the results of the placement tests, the students in the 2021 summer program were placed into one of the four initial curricular levels¹² (the levels are described earlier in this chapter).

As I was given access to the program's internal learning management system, I was able to explore the items on the program's placement and exit lexicogrammar quizzes (for examples, see Figures 4 and 5). The program's FIB and MC lexicogrammar quizzes cover various

¹² The participants from the dissertation research pilot study (Pastushenkov, 2020) were placed into one of the seven curricular levels during the face-to-face program of 2019.

grammatical and lexical forms. The tests cover a wide range of grammar rules from novice level structures such as the prepositional case of nouns to more advanced grammatical patterns such as the Russian verbs of motion. The FIB quiz includes 100 questions; the answers must be in Russian. For example, to successfully complete Question 1 in Figure 4, the test takers will need to have considerable knowledge of Russian vocabulary and various grammatical patterns such as tenses, comparative and superlative adjectives, passive voice, and pronoun-adjective-noun agreement in grammatical gender, number, and case. The students received scores on a 0-100 scale; these results were used in the analysis for RQ 2 as described in Chapter 7.

O родственниках (1. of [my] father)	я знаю (2. more than)	
маминых. Дед (3. was born)	в Петербурге и работал (4. [as a] tea	cher)
. Семья была (5. the s	implest) : ну, учитель и	
учительница. Но дед, (6. obviously)	, ещё хорошо готовил, (7. bea	cause)
(8. after the war)	ero (9. was invited)	
работу в подмосковные Горки, где (10. liv	ved) Ленин.	
Когда Ленин (11. died)	, (12. [my] grandfather)	перевели
на одну (13. of the summer houses)	Сталина, и он там (14. for a lo	ong time)

Figure 4. Sample question from the FIB lexicogrammar quiz.

The lexicogrammar MC quiz includes 150 questions. While completing the quiz, test takers are discouraged from using wild or strategic guessing as indicated in the program's

internal learning management system. These instructions are given to the students to make placement decisions more accurate. Similarly to the lexicogrammar FIB quiz, the MC quiz covers a wide range of grammatical and lexical forms. For example, Question 2 in Figure 5 tests L2 Russian learners' knowledge of noun-adjective agreement in the accusative case and Question 3 focuses on their knowledge of verb conjugation. In the final section of the MC lexicogrammar quiz, test takers are given four reading passages with multiple spaces that are left blank. Future students in the summer program need to read the passages first, and then answer a series of questions. The passages include excerpts from Russian literature, which is consistent with the fact that the summer program puts an emphasis on literature (especially at higher proficiency levels). The students received scores on a 0-150 scale; these results were used in the analysis for RQ 2 as described in Chapter 7.

D	Question 2	1 pts
	Мы слушаем	
	 русскую музыку 	
	о русской музыки	
	о русской музыке	
	о русская музыка	
l		
	Question 3	1 pts
	Отец часто письма.	
	⊙ пишут	
	О пишет	
	О пишем	
	⊙ пишете	

Figure 5. Sample questions from the MC lexicogrammar quiz.

WM tasks

In line with previous research (e.g., Serafini, 2017; Serafini & Sanz, 2016), this study adopted two complex measures of WM capacity conducted in English: an OSpan used to tap into the executive component of WM (the sequence is shown in Figure 6) and a Backward Digit Span (BDS) task measuring the participants' WM storage capacity (the sequence is shown Figure 7). The full list of stimuli (equations and English words for the OSpan task and sequences of numbers for the BDS task) is available on the researcher's website. In both tasks, the participants completed a practice round first. The tests were adopted from the library of experiments in Gorilla Experiment Builder, an online platform for behavioral scientists. In the recruitment email for Phase 1 when the WM data were collected, I mentioned that completing the tests on their laptops or personal computers may be easier for some of the participants even though they could also complete the tests on their smartphones. In the OSpan task, the participants saw a series of equations on the screen (e.g., $(9 + 5) \ge 1 = 15$). They needed to press T if the equation was true and F if the equation was false. The participants were asked to respond as quickly as possible. After each equation, they saw an English word (e.g., log or farm). The students were instructed to remember these words in the order that they saw them and type the words one at a time when prompted, pressing enter between the words. The participants were encouraged to guess the right word if they were unsure. The version of the OSpan task that I used included 12 sets of equations and English words. Each set ranged from two to six equations, followed by English words. The participants saw 50 equations total (processing component) and 50 English words total (storage component). On average, the OSpan task took approximately 8 minutes to complete. The scoring procedure for this WM test is described in Chapter 5; the procedure considers both the processing and storage components of OSpan.



Figure 6. Sequence of the Operation Span Task.

In the version of the BDS task that I used, the participants saw a sequence of digits (from 2 to 9 digits). Using the digit pad on the screen in the Gorilla interface, the students were required to repeat the digits in reverse. If a student failed to correctly repeat two sequences in a row, the BDS task was concluded. This task took 8 to 9 minutes to complete. The scoring procedure for the backward digit span task is described in Chapter 5.



Figure 7. Sequence of the Backward Digit Span Task.

Motivation surveys: L2MSS and General Motivation

To assess L2 Russian learners' motivation and to investigate how this ID changed over the course of the summer immersion program, I used two motivation surveys (referred to as L2MSS survey and General Motivation survey in this dissertation). The surveys were distributed at different phases of the study using Qualtrics, a cloud-based service for creating online surveys. The survey items (20 questions in the L2MSS survey and five questions in the General Motivation survey) are shown in Appendices A and B. When designing the surveys, I considered my experience with the dissertation pilot study (Pastushenkov, 2020) and my frequent consultations with the program director and faculty (for more detail, see Chapter 5). After visiting the program in 2019, I learned that it was difficult for some of the participants to find time to participate in the research. Considering that the surveys were distributed with the other ID measures (WM tests or additional L2 exposure logs) during several phases of the study, the motivation surveys were short, which is also important from pedagogical considerations discussed in Chapter 8. The L2MSS survey usually took up to 3-4 minutes to complete and the General Motivation survey took up to 1 minute. Considering that the number of items included in surveys was lower than in the motivation instruments used in previous research (e.g., Serafini, 2017), I assessed the surveys' reliability and construct validity to ensure that they were reliable tools that were also indicative of the students' motivation (for more information about reliability and validity analyses, see Chapter 5).

In addition to the three traditional components (ideal L2 self, ought-to L2 self, and learning experience) used in Dörnyei's (2009) L2MSS, I added the fourth component called antiought-to L2 self (see Pastushenkov & McIntyre, 2020; Thompson, 2017, 2021). To assess each component of the students' L2MSS, I used five questions (20 questions total in the L2MSS survey). An example of how the L2MSS survey items were shown to the participants in Qualtrics is given in Figure 8. The L2MSS questions were adapted from the studies by Thompson (2017) and Pastushenkov and McIntyre (2020). To obtain a more fine-grained understanding of motivation, the participants rated their motivation assessed by the L2MSS survey on a 10-point (0-9) Likert scale (on a continuum from strongly disagree to strongly agree)

70

Please answer these questions about your motivation to study Russian.

	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9 (Strongly agree)
l can imagine myself as someone who is able to speak Russian.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l can imagine speaking Russian as if I were a native speaker.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I can imagine myself speaking Russian with international colleagues.	\bigcirc	0	\bigcirc							
l can imagine myself living abroad and having a discussion in Russian.	\bigcirc	0	\bigcirc	\bigcirc	0	0	\bigcirc	0	\bigcirc	\bigcirc
I can imagine myself studying in a university where all my courses are taught in Russian.	0	\bigcirc	0	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	0	\bigcirc

Figure 8. Example of L2MSS survey items in Qualtrics.

The 10-point (0-9) Likert scale was also used in the General Motivation survey (see Figure 8 or Appendix B). The questions were adapted from the Motivation subset (effort, desire, and attitudes toward learning) of the Attitudes/Motivation Test Battery (AMTB) developed by Gardner and colleagues (e.g., Gardner et al., 2004). This test battery was also used by Serafini (2017) in addition to the L2MSS survey. The General Motivation survey included five items, focusing specifically on the learners' current motivation to study Russian. Rate your CURRENT motivation to learn Russian.

	0 (Very low)	1	2	3	4	5	6	7	8	9 (Very high)
My motivation to learn Russian in order to communicate with Russian speaking people is	\bigcirc	0	0	0	0	0	0	0	\bigcirc	\bigcirc
My desire to learn Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
My motivation to learn Russian for practical purposes (e.g., to get a job) is	0	\bigcirc	0	\bigcirc	0	0	0	0	0	\bigcirc
My interest in Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
My motivation to learn Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc

Figure 9. General Motivation survey in Qualtrics.

Additional L2 exposure log

In addition to meeting for at least 15 synchronous hours per week for eight weeks in the summer of 2021 (approximately 120 hours of instruction), the students were also engaged in learning Russian outside of their formal classes. To assess how students learned Russian outside of formal classes over the course of the summer program, I prepared an additional L2 exposure log distributed via Qualtrics (see Figure 10 or Appendix C). The log was distributed together with the motivation survey(s) during several phases of the study. The items on the additional L2 exposure log were adapted from Ranta and Meckelborg's (2013) language exposure log and included questions about the students' social interactions, academic work, and recreation. The

log was prepared based on the pilot study (Pastushenkov, 2020), my consultations with the program director and faculty, and the analysis of the students' schedule in the summer program. The goal was to align the items on the additional L2 exposure log with the program's cocurricular activities (e.g., informal conversations with peers and tutoring sessions), as well as with the common activities that students were known to be engaged in (e.g., watching movies in Russian, listening to Russian music, and using language learning applications [apps]). In the dissertation pilot study (Pastushenkov, 2020), I included several "Other" options in the log, but the participants generally skipped these options; therefore, I decided to update the original log and include more options for the students. The updated additional L2 exposure log included 13 items (see Figure 9). The students self-reported the total hours per week for Weeks 2, 4, 6, and 8 corresponding to Phases 2, 3, 4, and 5 of this study. Based on my consultation with the program director and faculty, I decided not to distribute the log during Weeks 1, 3, 5, and 7, as the number of phases in the study was already high (the students also tended to be busier during Week 1 of the summer program). The participants were instructed to report only the hours that they spent on activities outside of their formal classes in the summer program.

This week at

(Monday, August 2 - Friday, August 6), how much time did

you spend in the following activities OUTSIDE formal classes?

	Hours this week
Online communication through social media/email using Russian	
Conversations with instructors in Russian	
Conversations with tutors in Russian	
Conversations with peers in Russian	
Homework in the Russian program	
Language learning apps to study Russian	
Using Russian while playing sports or video games	
Reading Russian literature	
Reading Russian news/blogs online	
Watching TV/movies/streaming services in Russian	
Listening to music in Russian	
Using Russian while participating in thematic clubs	
Attending Russian guest lectures	

Figure 10. Additional L2 exposure log in Qualtrics. *Note.* The name of the summer program was deleted in this figure.

Post-program interviews

The post-program interviews were semi-structured and recorded in Zoom during the Fall 2021 semester after the students completed the summer program. Each interview took approximately 20 minutes. My goal was to recruit participants from Curricular Levels 1, 2, 3, and 4. The main interview questions aligned with RQs 1, 2, and are shown in Appendix D. During the interviews, I asked the participants about the changes in their motivation and

additional L2 exposure over the course of the summer (aligned with RQ 1) and about how WM, motivation, and additional L2 exposure affected their learning gains (aligned with RQ 2). For exploratory and pedagogical purposes, I also asked participants about what sources of additional L2 exposure (movies, music, apps, etc.) were particularly helpful for them. If necessary, I asked participants clarification questions and explained some of the concepts that we talked about.

The questions about motivation were aligned with the L2MSS and General Motivation surveys. I asked the interviewees to explain to what extent their general motivation to study Russian changed over the course of the summer programs, which aligned with the General Motivation survey. I also asked interviewees about whether they envisioned themselves as someone who will speak Russian in the future and whether this vision changed over the course of the summer program (aligned with ideal L2 self). I asked participants about any external pressures (e.g., from parents or teachers) that the students experienced when they chose to study Russian and whether these external pressures became weaker or stronger during their studies (aligned with ought-to L2 self). In another question related to motivation, I asked interviewees about whether anyone encouraged them not to study Russian (aligned with anti-ought-to L2 self). In the motivation part of the interview, I also asked students to describe their Russian learning experience in the summer program and at their previous institutions (aligned with the learning experience component of L2MSS).

The recorded interviews were transcribed and coded. The coding procedure followed previous qualitative research with L2 Russian learners, specifically the study by Zaykovskaya et al., 2017. In line with this research, I followed the principle of trusting the interviewees and avoided cleaning the qualitative data. The participants' statements were used at face value and any inconsistencies in the interviewees' statements were considered "signals of authenticity"

75

(Zaykovskaya et al., 2017, p. 116). As the main interview questions were aligned with the study's RQs and focused on one ID at a time (except for the last question about the relationships between IDs), the data coding procedure was relatively straightforward and included two coding categories. The first category was whether the interview data referred to RQ 1 (A and B) or RQ 2. The second coding category was based on the IDs (or ID facets) that the participants talked about: general motivation, ideal L2 self, ought-to L2 self, anti-ought-to L2 self, learning experience, additional L2 exposure, or WM.

Procedure

The procedure is shown in Figure 11. The study was conducted entirely online using Qualtrics for the consent form, background questionnaire, the L2MSS and General Motivation surveys, and additional L2 exposure log; Gorilla Experiment Builder for the WM tests; and Zoom for the interviews. This mixed-methods study included five main phases, during which different ID data were collected, followed by post-program interviews with the students. The ID data were collected during the summer of 2021; the qualitative interview data were collected during the Fall 2021 semester.

As cognitive abilities are considered more stable constructs (Serafini, 2017) than motivation (Boo et al, 2015) and exposure (Ranta & Meckelborg, 2013), the WM tests were administered once prior to the program. As L2MSS (particularly, the ideal L2 self) has been considered a more stable construct than General Motivation (see Pastushenkov & McIntyre, 2020; Serafini, 2017), the L2MSS data were collected twice (during Phase 1 and Phase 5), whereas the General Motivation data were collected during each phase of the study. The participants reported on their additional L2 Russian exposure during activities outside of their formal classes for Weeks 2, 4, 6, and 8, corresponding to Phases 2, 3, 4, and 5 of the study. The

76

participants were sent multiple emails about the study (for Phases 2 through 5, the emails were sent on Fridays). For each phase of the study, the participants were also given completion deadlines (Sundays of Weeks 2, 4, 6, and 8). The summaries of analyses used to address RQs 1 and 2 are discussed in the beginning of Chapters 6 and 7. The other important methodological considerations are discussed in Chapter 4 (pre-study preparation) and Chapter 5 (data processing).

Phase 1:	Phase 2:	Phase 3:	Phase 4:	Phase 5:	Interviews:
before Week 1	Week 2	Week 4	Week 6	Week 8	after Week 8
15-20 min.	5-10 min.	5-10 min.	5-10 min.	5-15 min.	15-20 min.
L2MSS GMS WM Tests	GMS Exp. log	GMS Exp. log	GMS Exp. log	L2MSS GMS Exp. Log	WM motivation Exp.
	Phase 1: before Week 1 15-20 min. L2MSS GMS WM Tests	Phase 1:Phase 2:before Week 1Week 215-20 min.5-10 min.L2MSSGMSGMSExp. logWM TestsS	Phase 1:Phase 2:Phase 3:before Week 1Week 2Week 415-20 min.5-10 min.5-10 min.L2MSSGMSGMSGMSExp. logExp. logWM TestsKK	Phase 1:Phase 2:Phase 3:Phase 4:before Week 1Week 2Week 4Week 615-20 min.5-10 min.5-10 min.5-10 min.L2MSSGMSGMSGMSGMSExp. logExp. logExp. logWM TestsKKK	Phase 1:Phase 2:Phase 3:Phase 4:Phase 5:before Week 1Week 2Week 4Week 6Week 815-20 min.5-10 min.5-10 min.5-10 min.5-15 min.L2MSSGMSGMSGMSL2MSSGMSExp. logExp. logExp. logGMSWM TestsFree Free Free Free Free Free Free Free

Figure 11. Procedure.

Notes. Week 8 = last week of the summer program; WM Tests = backward digit span (BDS) task and operation span task (OSpan); L2MSS = L2 Motivational Self System, GMS = General Motivation survey; Exp. log = additional L2 exposure log.

Summary of Chapter 3

This chapter provided a thorough description of the eight-week Russian summer immersion program, where the data for this dissertation research were collected. Due to the fact that the data were collected during the COVID-19 pandemic, the program included four curricular levels instead of seven. Each curricular level is discussed in detail, including the number of hours of prior Russian exposure that the students were expected to have and the program's experience with regards to the students' progress, based on the gains on the ACTFL scale. In this chapter, I also discussed some of the distinct features of the summer program such as the language pledge that the students need to take prior to starting their studies, the program's emphasis on Russian culture and individualized approach, and that the 2021 program was valued at 9 credit hours and participants met for at least 15 synchronous hours per week; this gave the students approximately 120 hours of instruction from the end of June to mid-August. I also described the background of the participants in the study (N = 52 out of 111 students in the summer program), who were considered representative of the students' population in the program. The large portion of this chapter focuses on the description of the materials that I used in my dissertation: the program's placement and exit tests (a lexicogrammar FIB quiz, a lexicogrammar MC quiz, ACTFL OPIc, and ACTFL WPT); the two WM tests that I used (OSpan and BDS); the two motivation surveys (L2MSS and General Motivation); and the additional L2 exposure log that was used to assess the participants' L2 Russian learning outside of their formal classes in the summer program. I also talked about the post-program interviews and the coding procedure that I used for the qualitative data. Finally, I discussed the procedure of this mixed-methods study that included multiple phases and was conducted over the course of several months.

CHAPTER 4: PRE-STUDY PREPARATION

Internal reviews

The present dissertation research was reviewed by the Internal Review Boards (IRB) at my home institution (Michigan State University) and at the summer program (the name of the summer program is kept anonymous throughout the manuscript). The appropriate administrator in the summer program provided approval for the use of protected student education records at their institution for research. To satisfy Family Educational Rights and Privacy Act (FERPA) regulations, the informed consent form included the following: a complete list of education records that I collected for research (the results of the placement and exit tests in the summer program); a statement regarding the purpose of the disclosure of this information for research (my goal was to explore the relationships between learners' IDs and their learning gains, for which the results of the placement and exit tests were used); and information about the researcher to whom the disclosure of research records is made. As the materials and procedure were modified after the pilot studies discussed in this chapter, I made modifications in the IRB application form that required additional review at my home institution and the summer program. The IRB boards reviewed the modifications and approved the study (see Appendix E).

Pilot study #1: Test equating

I conducted a separate pilot study (my second Qualifying Research Paper in the Ph.D. Program in Second Language Studies at Michigan State University) to verify that the two program-administered lexicogrammar tests¹³ were equally difficult and to equate the tests if

¹³ Since I did not have access to the speaking (ACTFL OPIs/OPIcs) and writing tests (ACTFL WPTs), no separate test equating procedures were carried out for these tests. The summer program strives to use equally difficult speaking and writing tests to reliably measure learning gains.

necessary (for more information about test equating, see LaFlair et al., 2017). The procedure was difficult to implement since test equating generally requires a large number of participants. Russian is a LCTL and the number of students in traditional Russian programs is relatively low compared to more commonly taught languages such as English and Spanish. Another difficulty was that the program's placement and exit lexicogrammar tests included a total of 500 items, making it impractical to distribute the placement and exit tests simultaneously¹⁴. Considering these circumstances, the program's FIB and MC quizzes were distributed separately in Russian programs at two universities with the help of my colleagues.

The entrance and exit FIB quizzes (200 items total; 100 questions each) were distributed during a Saturday session to a group of five students who at that time had completed the summer immersion component of a Russian immersion program and were in their second semester of the follow-up academic year. The students were given 100 minutes to complete the tests. The items were presented in random order. Based on my colleague's evaluation of, the learners demonstrated a 'floor' effect, with the mean scores for both the placement and exit tests barely reaching 10%. Considering my colleague's evaluation and the fact that the exit and placement FIB quizzes focused on similar lexicogrammatical patterns (for more information, see Chapter 3), my assumption was that the FIB quizzes were approximately equally difficult and could be used in measuring learning gains in the summer program.

A different strategy was used to fit the MC quizzes into 50-minute lessons at another university. Upon my consultation with a dissertation committee member, eleven versions of combined placement and exit MC tests with 50 items and 26 anchors each (the same items used

¹⁴ Organizations such as Educational Testing Service have the capacity to distribute different versions of their standardized tests simultaneously during the exam sessions to ensure that their tests are equally difficult.

in each version of the test) were prepared. The participants were 33 intermediate students (3 students*11 versions) who received extra credit for their participation in the pilot study. The mean scores and SDs for the placement and exit tests were calculated. The scores were approximately equal for both placement test anchors (M = 0.46, SD = 0.17) and exit test anchors (M = 0.50, SD = 0.14). Considering these findings and the fact that the placement and exit MC tests followed a similar structure and focused on similar lexicogrammatical structures (see Chapter 3 for more detail), my assumption was that the tests were equally difficult. Therefore, I did not conduct additional statistical test equating. Due the lack of statistical power, which is considered a limitation of this dissertation research, the results for the learning gains on the program's lexicogrammar quizzes should be interpreted with caution.

Pilot study #2: Dissertation pilot study

To test the procedure's feasibility, a pilot study (Pastushenkov, 2020) was conducted in the same Russian summer immersion program. The pilot study focused on language learning apps¹⁵, a source of additional L2 exposure that some of the students in the summer program were engaged in outside their formal classes. The pilot study aimed to investigate the role of languagelearning apps in L2 grammar learning gains as measured by the program's placement and exit MC grammar quiz, as well as test various instruments and procedures used in this dissertation research. In the pilot study, I also explored the relationships between L2 students' use of language learning apps, their motivation as assessed by an L2MSS survey, and other sources of L2 exposure assessed by utilizing a simplified version of the additional L2 exposure log used in the dissertation research. The data were collected in the face-to-face program during the summer

¹⁵ App use was also investigated in the present dissertation research along with the other sources of additional L2 exposure.

of 2019. The procedure included one phase during which the quantitative ID data (WM, additional L2 exposure, and motivation) were collected. The WM data were not included in the 2020 publication as a part of the special issue on "Digital Culture and Humanities," as the topic of WM was beyond the scope of the issue. As it was done later in my dissertation research, the participants in the pilot study self-reported their language learning app use in mean hours per week before and during the summer program. The participants were 27 students who were recruited from seven curricular levels: Level 1 Introductory Russian (n = 3), Level 2 Advanced Intermediate Russian (n = 4), Level 3 Basic Intermediate Russian (n = 1), Level 4 Enhanced Russian I (n = 7), and Level 7 Advanced Russian II (n = 3).

The data were analyzed for all levels combined and for the beginner/intermediate and advanced levels separately (curricular levels 1-4 and 5-7 were merged into two levels due to the low sample size). The multiple regression analysis revealed a small effect size of language learning app use on grammar learning gains as measured by the MC quiz during the program ($R^2 = .12$, p = .31) for all curricular levels combined (levels 1-7). However, a large effect size was found for levels 1-4 ($R^2 = .43$, p = .23), suggesting that language learning apps were a helpful addition to the Russian learners' daily routines in the summer program (or a part of daily routines of high achievers), particularly at the beginner and intermediate levels. Moreover, a statistically significant positive correlation between the students' app use for all levels combined before and during the summer program was found (r = .58, p < 0.01), suggesting that the participants used language learning apps consistently during their Russian language studies, including the summer program and their prior institution(s). In terms of motivation, the pilot study did not reveal any statistically or practically significant correlations between app use before or during the program

and any aspects of L2MSS for the curricular levels 1-7, 1-4, and 5-7. It is important to note that the pilot study included a low sample size, and the regression models and correlation analyses were underpowered and thus the results need to be interpreted with caution.

Dissertation research modifications

Despite its limitations, particularly the limited sample size and the fact that I had to merge the results for the beginner and intermediate participants for some parts of the analyses, the pilot study helped identify limitations and future directions and update the materials and procedure of my dissertation. The key differences between the pilot study and the present dissertation research are shown in Table 2. First and foremost, I aimed to expand the sample size in my dissertation and investigate the role of IDs at all levels combined and at each curricular level separately. Another major limitation of the pilot study was that it included only one data point. IDs such as motivation and L2 exposure are generally considered as non-static, dynamic constructs, which also aligns with DST that views L2 learning as a co-regulated process. Therefore, collecting data at multiple points would have been beneficial. This limitation was addressed in my dissertation, which included five phases during which quantitative ID data were collected through the motivation surveys and additional L2 exposure log data collected during several phases (the WM data were collected once). The updated procedure also included postprogram interviews with the participants when the qualitative data were collected. The switch to the mixed-methods design was made to gain a deeper understanding of the role of IDs in L2 Russian learning.

83

	Pilot Study (Pastushenkov, 2020)	Dissertation Research
Data collection	July-August, 2019	June-November, 2021
Instruction mode	Face-to-face	Online
Curricular levels	7	4
N of participants	27	52
Types of data	Quantitative	Mixed
Procedure	One phase	Five phases and post-
		program interviews
Placement/exit	Lexicogrammar MC quiz	Lexicogrammar MC quiz,
tests		lexicogrammar FIB quiz,
		ACTFL OPIcs, and ACTFL
		WPTs
WM tests	BDS	OSpan and BDS
L2 exposure log	A simplified version of the additional L2	An extended version of the
	exposure log	additional L2 exposure log
Motivation	L2MSS	L2MSS and General
surveys		Motivation

Table 2. Key differences between the dissertation pilot study and dissertation research

Through the pilot study, I learned that it was difficult for some of the students in the summer program to find time to participate in the research study. To address this issue, I collected WM data, which was the most time-consuming phase of the study, right before the program. In line with the works by Serafini and colleagues, WM data were collected once, as this cognitive ID is considered a more stable construct than motivation and L2 exposure. Following the research of Serafini (2017), I also added a second L2 motivational framework (General Motivation) to complement the L2MSS used in the pilot study. In this dissertation research, the General Motivation survey was adapted from the Motivation subset of the AMTB test battery; the questions were simplified and focused on the participants' general motivation to study Russian since it was a focus on the study, and it was important not to use lengthy surveys during the summer program when the participants were busy. Additional L2 exposure log was also modified based upon my visit to the summer program in 2019 and upon my consultations with the summer program's faculty and students. The log was extended to obtain a better view of how

students learned Russian outside their formal classes and included topics that were specific to the summer program (e.g., conversations with tutors in Russian or using Russian while participating in thematic clubs).

The BDS task measuring the participants' WM storage capacity was used in the pilot study (these results were not included in Pastushenkov, 2020). No long-term WM effects on the grammar learning gains as measured by the MC test were found. Moreover, there was no evidence of any relationships between WM capacity and L2 exposure (i.e., the lack of WM capacity did not limit the students' exposure to Russian outside their formal classes as it was originally hypothesized). In line with Serafini and colleagues, OSpan, another measure of WM, was added to the dissertation research to tap into the executive component of WM and obtain a better understanding of the role of WM in L2 Russian learning.

In this dissertation research, I also added several dependent variables. In addition to the learning gains on the MC quiz, I added the second grammar quiz (FIB), along with the speaking¹⁶ and writing learning gains. Upon my consultation with the program director and faculty and my collaboration with my co-authors on Merrill et al. (2021) who had considerable experience working in Russian immersion programs, the grammar quizzes (the terminology that I used in the pilot study) are referred to lexicogrammar quizzes in this dissertation to better reflect the content of the quizzes.

Upon completion of the pilot study (Pastushenkov, 2020) in the face-to-face program of 2019, my plan was to return to the program in the summer of 2020. Due to the COVID-19 pandemic and the switch to emergency remote instruction in the summer program, I decided to

¹⁶ The online program of 2021 used computerized ACTFL OPIcs, whereas the 2019 program used 'live' OPIs.

postpone my dissertation data collection until the summer of 2021. Although the 2021 program was also conducted online, this time it was not considered emergency remote instruction. More L2 Russian students signed up compared to the summer of 2020; moreover, the program was better prepared for online teaching, as they had more experience in this area¹⁷. One of the key differences between the face-to-face program of 2019 and the online program of 2021 was the number of curricular levels. The face-to-face program included seven levels, whereas the online program included four levels¹⁸. This feature of the online program was taken into consideration in the analyses.

Summary of Chapter 4

The present dissertation research and its modifications that were made upon completion of the two pilot studies were reviewed and approved by the IRBs at my home institution (Michigan State University) and at the summer program (see Appendix E). The first pilot study described in this chapter was conducted to verify that the two program-administered lexicogrammar tests were equally difficult and if test equating was necessary. I did not conduct a separate pilot study for the ACTFL OPIs/OPIcs and WPTs because I did not have access to the test materials. Even though the pilot study was difficult to implement and had a number of limitations, it was possible to make an assumption that the two lexicogrammar tests were equally difficult and could be used in calculating the learning gains. Since ACTFL and the summer program ensures the use of equally difficult speaking and writing tests, the ACTFL OPIcs and WPTs could also be used in calculating the gains. To test the procedure's feasibility for dissertation research, the second pilot study was conducted (Pastushenkov, 2020). Despite its

¹⁷ The key aspects of the program's online pedagogy, including the number of contact hours, are discussed in detail in Chapter 3.

¹⁸ The program and levels are described in detail in Chapter 3.

limitations (specifically the limited sample size and the inclusion of only one phase in the data collection), this pilot study helped identify a number of areas that needed modifications in the dissertation research. The key differences between the studies included the recruitment and number of participants in different curricular levels, the number of phases in the procedure, the need for the switch to the mixed methods methodology, the focus on several types of learning gains, the addition of Ospan as the second WM test, the revisions of the additional L2 exposure log based on the activities that the students in the summer program are engaged in, and the addition of a second motivational framework in line with previous research.

CHAPTER 5: DATA PROCESSING

Data collection

The quantitative ID data were collected online during five phases in the summer of 2021 online via Qualtrics and Gorilla Experiment Builder: one week before the program and on a biweekly basis during the summer program. Different data were collected during each phase (see Figure 11 for the procedure). The raw ID data were exported from Qualtrics and Gorilla Experiment Builder and then underwent initial preparation and data cleaning procedures discussed in detail in this chapter. After the background questionnaire and motivation surveys during Phase 1, participants were transferred to the WM tests in Gorilla. The participants were assigned an ID in Qualtrics; this ID helped connect the data from Qualtrics and Gorilla. Upon approval from the IRBs at my home institution and the summer program (see Appendix E), I was then granted access to the testing data collected by the summer program, which were used to calculate learning gains. The interview data were collected online via Zoom after the completion of the summer program in the fall of 2021.

Data preparation

Initial data preparation

The raw WM data for the OSpan and BDS tasks from Gorilla Experiment Builder underwent initial cleaning. Only complete responses¹⁹ from the WM tests were used in the analyses. Complete responses from 42 out of 52 participants on Phase 1 were collected. Three participants completed the tests twice; only their first attempt was used in the analyses. Due to potential technical problems, the BDS task took over an hour for one of the participants. This

¹⁹ Gorilla Experiment Builder marks responses as "complete" and "live" ("live" responses were excluded during initial data cleaning).

value was excluded from further analyses. The mean completion time of the OSpan task was 503.25 seconds (SD = 151.80) and for the BDS task 518.16 seconds (SD = 316.31). In line with the main data cleaning procedure discussed later in this chapter, one response beyond three SDs from the mean (1173 seconds on the OSpan task) was excluded from further analyses.

The scores for the BDS task were automatically calculated in Gorilla. Two participants who scored 0 on the BDS task were excluded from further analysis (their OSpan scores were also discarded). The scores for processing (0-50 scale) and storage (0-50 scale) were not automatically calculated in the version of the OSpan task that I used. SLA researchers have previously used different methods for calculating scores on WM tests (e.g., Absolute, Total, and Lenient for the storage component of the OSpan task, see Serafini & Sanz, 2016). For the purpose of simplicity and given the inter-correlations between the methods (see Serafini, 2017), the Total scoring method was used for both the processing and storage components using the participants' responses recorded in Gorilla: true or false for the 50 equations total (processing component) and for the 50 English words total that the participants saw after each equation and then typed after each set of equations²⁰ (storage component). Only correct responses for the equations (e.g., TRUE for $(2 + 1) \times 5 = 15$) counted toward the processing score. Only complete responses with correctly spelled English words written in the order that they were presented after each equation counted toward the storage score. Prior to the main data cleaning based on standard deviations, the participants who scored 0 on processing or storage were excluded (two responses total). After calculating the processing and storage scores, the composite total OSpan task score was calculated (0-100 scale) and used in the analyses.

²⁰ The version of the OSpan task that I used included 12 sets of equations/English words. Each set of equations ranged from two to six equations followed by English words.

The raw motivation data were exported from Qualtrics. During the initial data cleaning, the responses on the two motivation surveys and L2 exposure log below 85% completion rate were excluded from further analyses in line with previous survey-based research (see Sato et al., 2021). The students were instructed to skip the question about their prior L2 Russian experience (part of the L2MSS survey) on Phase 1 if they had not studied Russian prior to the summer program and joined the program as absolute beginners (if the participants skipped this question, it did not count toward the 85% completion threshold). A total of three responses were excluded during the initial cleaning of motivation data.

The L2MSS data included four components and five questions (referred to as items) in each component (ideal L2 self, ought-to L2 self, anti-ought-to L2 self, and learning experience) and General Motivation that included one section with five items that focused on students' general motivation to study Russian. Mean scores were calculated for each component of the L2MSS survey for Phases 1 and 5 and for the General Motivation survey for Phases 2, 3, 4, and 5. The participants completed additional L2 exposure logs during Phases 2, 3, 4, and 5 and reported how much time they spent on 13 different activities to learn Russian. Total additional L2 exposure in the 13 categories were calculated for Phases 2-5. Additionally, I calculated the number of hours in each category²¹ to explore what types of additional L2 exposure the participants engaged in.

The learning gains were calculated by subtracting the students' placement scores from their exit scores on the lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and

²¹ The number of hours in each separate category of additional L2 exposure was added to this dissertation for exploratory and pedagogical purposes.

ACTFL WPTs. For the purposes of the analyses, the scores on the ACTFL OPIcs and WPTs were converted to a numeric scale (for more information, see Chapter 3).

Data cleaning, missing data, and outliers

The number of students from the Russian summer program who participated in different phases of the study varied: Phase 1 (n = 52), Phase 2 (n = 34), Phase 3 (n = 33), Phase 4 (n = 29), Phase 5 (n = 24), and post-program interviews (n = 11). These discrepancies resulted in missing data. To maximize the number of observations and minimize the data loss that occurred as a result of listwise deletion, pairwise deletion was used throughout the analyses. Despite its benefits, pairwise deletion may lead to biased parameter estimates (Lodder, 2013). To address this issue, researchers have used listwise deletion with larger samples (see Kisselev et al., forthcoming) and other methods of dealing with missing data (e.g., full-information maximum likelihood method, see Sainani, 2015). The discrepancies in the number of students who participated in different phases of the study, which resulted in missing data and subsequent use of pairwise deletion, are considered a limitation and are discussed in detail in Chapter 8.

For the purpose of simplicity and in order to maximize the number of observations, I followed a less demanding approach to data cleaning and analyzed the data for outliers beyond three standard deviations from the mean for the four curricular levels combined (any values beyond mean plus standard deviation multiplied by 3 and mean minus standard deviation multiplied by 3 for the independent and dependent variables). The results of the calculations are shown in Table 3. Despite its benefits, the mean plus or minus three has a number of disadvantages (e.g., the mean and standard deviations are affected by outliers, see Leys et al., 2013). Other methods for outlier analysis are discussed in Chapter 8.

A more liberal approach for data cleaning was used for the motivation surveys (L2MSS and General Motivation) and learning gains because of pedagogical considerations. For example, one of the participants indicated that their ideal L2 self in Phase 5 at the end of the summer was 0 (0 on all five items on this part of the L2MSS survey). Although this is an outlier from a statistical point of view, it is also a realistic response and a potentially important finding from a pedagogical standpoint. Considering that the motivation data had undergone initial cleaning described earlier in this chapter, I decided not to exclude the additional outliers in the motivation data (three values total) to maximize the number of observations. The same approach was used for the learning gains (one potential outlier was kept for further analyses). A more conservative approach was adopted for the WM tests and additional L2 exposure log. One outlier was excluded from the OSpan Composite score during this stage of data cleaning, as the extreme WM scores did not have direct pedagogical implications as opposed to the self-reported motivation survey data on a 0-9 scale. Moreover, six outliers were excluded from the additional L2 exposure data. From a pedagogical standpoint, these outliers were also unrealistically high and needed to be excluded (e.g., 94.50 hours of L2 exposure).

		M + 3*SD	M - 3*SD	# Outliers			
	INDEPENDE	NT VARIABLE	ES				
	P	hase 1					
WM	OSpan Composite	110.33	51.25	1 (excluded)			
	BDS	49.94	-6.14	0			
L2MSS	Ideal L2-self	12.13	0.42	0			
	Ought-to L2 self	8.73	-3.02	0			
	Anti-ought-to L2 self	10.77	-1.15	0			
	Learning experience	11.76	3.48	0			
General Motiva	tion	11.31	4.31	1 (not			
				excluded)			
	P	hase 2					
General Motiva	tion	11.32	3.85	0			
Additional L2 e	xposure	6.89	-1.77	2 (excluded)			
	P	hase 3					
General Motiva	tion	10.87	4.30	0			
Additional L2 e	xposure	6.62	-1.38	2 (excluded)			
	P	hase 4					
General Motiva	tion	11.87	2.86	1 (not			
				excluded)			
Additional L2 e	xposure	6.30	-1.67	1 (excluded)			
	P	hase 5					
L2MSS	Ideal L2-self	13.25	1.06	1 (not			
				excluded)			
	Ought-to L2 self	10.62	-3.44	0			
	Anti-ought-to L2 self	11.09	-1.88	0			
	Learning experience	12.18	3.47	0			
General Motiva	tion	12.44	3.16	1 (not			
				excluded)			
Additional L2 e	xposure	5.70	-1.62	1 (excluded)			
DEPENDENT VARIABLES							
Lexicogrammar	FIB	54.45	-25.47	1 (not			
				excluded)			
	MC	80.16	-30.53	0			
Speaking (ACT	FL OPIcs)	6.47	-2.60	0			
Writing (ACTF	L WPTs)	7.75	-3.64	0			

Table 3. Outlier analysis (four curricular levels combined)

Reliability and validity

In line with previous research (e.g., Kisselev et al., forthcoming; Serafini, 2017), the Cronbach's alpha coefficients were used for reliability analysis. The reliability coefficients were calculated for the two motivation surveys only as the additional L2 exposure log included a nonunitary scale (see Table 4). Additionally, the reliability coefficients were calculated for the four groups of independent variables (WM tests, L2MSS, and General Motivation) and the four dependent variables (learning on the lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs). The coefficients for the dependent and independent variables are shown in Table 5. The interpretations of the reliability coefficients were based on previous research (e.g., Taber, 2018). The Cronbach's alpha coefficients for the L2MSS survey on Phases 1 (α = .84) and 5 (α = .91) and the General Motivation survey on Phases 1 (α = .80), 2 (α = .88), 3 (α = .88), 4 (α = .93), and 5 (α = .93) indicated fairly high levels of internal consistency.

comethea)			
		# Items	Cronbach's
			alpha
	Phase 1		
L2MSS survey		20	.84
General Motivation survey		5	.88
	Phase 2		
General Motivation survey		5	.88
	Phase 3		
General Motivation survey		5	.88
	Phase 4		
General Motivation survey		5	.93
	Phase 5		
L2MSS survey		20	.91
General Motivation survey		5	.93

 Table 4. Reliability analysis (motivation surveys; four curricular levels combined)

The Cronbach's alpha coefficients for the independent variables were the following: the OSpan and BDS WM tests for Phase 1 (two variables; $\alpha = .72$); the L2MSS survey for Phases 1 and 5 (10 variables including the two composite L2MSS scores; $\alpha = .78$); and the General Motivation survey for Phases 1-5 (five variables; $\alpha = .84$). These statistics were also calculated for the learning gains on the lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs (four dependent variables; $\alpha = .87$). The Cronbach's alpha coefficients

indicated relatively high levels of internal consistency for most of the independent and dependent variables with the exception of the WM tests, for which internal consistency was moderate but acceptable.

Table 5. Reliability analysis (dependent and independent variables; four curricular levels combined) Image: state of the state o						
	Phases	#	Cronbach's	Cronbach's		
		Variables	alpha	alpha based		
				on		
				standardized		
				items		
WM tests (OSpan Composite and BDS)	1	2		.72		
L2MSS survey	1, 5	10	.78			
General Motivation survey	1, 2, 3, 4, 5	5	.84			
Learning gains	N/A	4		.87		

Since the sample size was relatively low for a factor analysis (particularly on Phases 2, 3, and 4), Pearson's correlation coefficients were calculated (for review, see Odom & Morrow, 2009) to assess construct validity of the instruments (i.e., how well the scores on the motivation surveys and responses for the additional L2 exposure log were indicative of the Russian language learners' IDs). Considering that the two WM tests had been previously validated and commonly used in psycholinguistic SLA research (e.g., Serafini, 2017), I did not conduct separate construct validity analysis for the OSpan and BDS tasks. I also did not assess the construct validity of the placement and exit lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs because this would require access to the responses on each individual item on the tests. Even though the summer program had successfully used these tests throughout its history, an investigation of the tests' validity would be beneficial and is considered a possible future direction for this project.
First, the total values for the groups of independent variables on Phases 1-5 were computed. Then the correlations between the total values and the items in each group of independent variables were calculated (the calculations along with the instruments will be available on the researcher's website). The obtained Pearson's r correlations were then compared against the critical values for r at a .05 significance level, taking the degrees of freedom (the total number of score pairs minus two) into consideration. The following instruments were validated: L2MSS survey on Phases 1 and 2 (no misfitting items were found) and General Motivation survey on Phase 1, 2, 3, 4, and 5 (no misfitting items were found). Therefore, the two motivation

Data input

Quantitative data

The independent variables included three main categories: WM, motivation (with L2MSS and General Motivation), and additional L2 exposure (referred to as the four groups of independent variables). Each category included several variables depending on the number of phases during which these ID data were collected. The independent variables included the following: two WM variables (BDS and OSpan*one phase); eight L2MSS survey variables (ideal L2 self, ought-to L2 self, anti-ought-to L2 self, and learning experience); five General Motivation variables (General Motivation*five phases); and four additional L2 exposure variables (additional L2 exposure*four phases). Composite WM and L2MSS scores were calculated for additional regression analyses (see Chapter 7). Since the data were explored for four curricular levels combined, separately for each curricular level, and individually for each participant, the independent variables included several sub-levels (all curricular levels combined; Level 1, Level 2, Level 3, and Level 4; and individual participant level). Four (main) dependent

96

variables were used in the analyses: learning gains on the lexicogrammar FIB quiz,

lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs (the placement and exit scores were used in additional regression analyses, see Chapter 7). The main independent and dependent variables used in the analyses are shown in Table 6.

Table 6. Main independent and	l dependent va	iriables use	d in the analyse	?S								
Variable	Phase	Scale	# Items	RQ								
INDEPENDENT VARIABLES												
	WM											
OSpan	1	0-100	50 processing and 50 storage	2								
BDS		0-50	N/A	2								
	MOTIVATI	ON										
Ideal L2-self Ought-to L2 self Anti-ought-to L2 self Learning experience	1,5	0-9	5	1, 2								
General Motivation	1, 2, 3, 4, 5											
ADDI	ΓΙΟΝΑL L2 Ε	EXPOSURI	Ξ									
Mean hours of additional L2 exposure per week for 13 categories combined	2, 3, 4, 5	0+	13	1, 2								
DEPI	ENDENT VA	RIABLES										
LEARNING GAINS												
Lexicogrammar FIB Lexicogrammar MC Speaking (ACTFL OPIcs) Writing (ACTFL WPTs)	N/A	0-100 0-150 0-10	100 150 12-17 4-5	1, 2								

Qualitative data

The qualitative data from the interviews with 11 participants were recorded via Zoom, transcribed, and then coded for RQs 1 and 2 and the ID (or ID facet) that the participants talked about (for more information, see Chapter 3). In line with previous qualitative research with L2 Russian learners (Zaykovskaya et al., 2017), I avoided cleaning the qualitative data. In line with methods used in previous literature on the qualitative methodologies (e.g., Eldh et al., 2020) and qualitative/mixed methods SLA research with L2 Russian learners (e.g., Pastushenkov & McIntyre, 2020; Zaykovskaya et al., 2017), quotations from the interviews were used in Chapters 6 and 7 to illustrate the results. Each interviewee was assigned an identification code (one English letter) to help distinguish the quotations from different participants. The quotations were complemented with a reference. The references included an excerpt number, an English letter representing the participant, their initial curricular level in the program (Level 1-4), and the ID(s) that the participants talked about in each excerpt. Here is a sample reference: [Excerpt X, Participant X, Level X, General Motivation]. The initial curricular level was included in the references for the quotations because the data were also analyzed separately for each level. Even though the interviews were conducted in English, the participants, especially those at higher proficiency levels, occasionally switched to Russian. The Russian segments were translated into English and bolded in the quotations.

Processing and output

The quantitative ID data and the testing data that were used to calculate learning gains were analyzed in IBM® SPSS® Statistics (version 27) and Microsoft Excel (2021 version). No specialized qualitative analysis software was used in this study. The interview data were coded in the transcripts (docx. files) using the comment feature in Microsoft Word. The results of the

98

quantitative data analysis are described in detail and represented visually in the form tables and graphs in Chapter 6 and 7. Building off the quantitative results, the results of the qualitative data analysis were added in Chapters 6, 7, and 8. First, the general trends found in the analysis of qualitative data were discussed (e.g., whether the qualitative data supported the quantitative analysis). The identified general trends were supported by quotations from the interviewees with individual references (for more information about the analyses and data triangulation, see analysis overviews in the beginning of Chapters 6 and 7).

Data storage

The data consisting of behavioral responses to the research instruments (two WM tests, two motivation surveys, additional L2 exposure log, four types of learning gains based on the summer program's placement and exit tests, and interview questions) are stored as .docx and .xlsx files on the researcher's personal password-protected computer and Google Docs and Spreadsheets shared with the PI. Participant identification codes were assigned to each participant to protect confidentiality and all files are stored with the codes. A tally sheet linking participant names and their identification numbers was kept separately in a secured .xlsx file on the researcher's personal password-protected computer. The participants' names are not disclosed in any of the shared documents. The tally sheet linking participant names and codes will not be shared and will be destroyed after completing the project. The instruments and anonymized data (including detailed meta-data) will be shared with the public through IRIS, a digital repository of instruments and materials for research into second languages (see Marsden et al., 2016).

99

Summary of Chapter 5

In Chapter 5, I discussed the following steps of data processing: data collection, data preparation (including the data cleaning and the analysis of reliability and validity), data input (quantitative and qualitative data), processing and output, and data storage. These steps were fundamental for this dissertation research and helped identify future steps in the analyses discussed in Chapters 6 and 7, as well as some of the limitations of the study (e.g., the relatively low reliability of the additional L2 exposure log). The quantitative data underwent several cleaning steps, and the potential outliers were excluded from further analyses. The instruments used in the data collection and the independent and dependent variables used in the analyses have shown acceptable levels of internal consistency, as measured by the Cronbach's alpha coefficients used in the reliability analysis. Based on the validity analysis, the two motivation surveys and additional L2 exposure log (with the exception of the additional L2 exposure log for Phases 4 and 5) were considered indicative of the participants' IDs. The clean data were prepared for further analyses.

CHAPTER 6: MOTIVATION AND L2 EXPOSURE AS DYNAMIC CONSTRUCTS Analysis overview of RQ 1

RQ 1 (A) focuses on the longitudinal changes in the participants' motivation assessed by the L2MSS and General Motivation surveys, and RQ 1 (B) on the amount of L2 exposure assessed by the bi-weekly log focusing on different Russian learning activities outside formal classes in the summer program (referred to as additional L2 exposure). To address RQ1, the quantitative data (descriptive statistics and 95% confidence intervals [CI] of the means for the ID data) and the qualitative data from the interviews (coded for RQ1 and focusing on motivation and additional L2 exposure²²) were analyzed for the four initial proficiency levels combined and for each curricular level separately. The descriptive statistics (including the number of students who participated in each phase of the study) and 95% CI for the quantitative ID data for Phases 1-5 are shown in Table 7 for the four levels combined and separately for Curricular Levels 1-4 in Tables 8, 9, and 10 with regards to WM, motivation, and L2 exposure, respectively²³. The general trends in the qualitative data regarding the changes in the participants' motivation and additional L2 exposure were described, including whether these trends aligned with the quantitative data. The main results identified in the analysis of the qualitative data were supported by quotations from the participants at different curricular levels.

To explore what Russian learning activities the participants were engaged in during the summer program, the total number of hours of different aspects of additional L2 exposure (e.g., communication via social media/email in Russian; conversations with instructors in Russian;

²² For more information about interview data coding, please see Chapter 3.

²³ The WM data are not discussed in this chapter since this type of cognitive ID is viewed as a more stable construct than motivation and L2 exposure (for review, see "WM in ISLA" in Chapter 3).

conversations with tutors in Russian, etc.) for Phases 2-5 were calculated for exploratory and pedagogical purposes. The descriptive statistics, along with 95% CI for the hours of different aspects of additional L2 exposure, are shown in Tables 11, 12, 13, and 14 for the four levels combined for Phases 2, 3, 4, and 5²⁴ and represented visually in Figure 26 for Phases 2-5.²⁵

The changes in the students' motivation during Phases 1-5 and their additional L2 Russian exposure during Phases 2-5 are also represented visually using box and whisker plots for the four proficiency levels combined and for Curricular Levels 1-4. In addition to the mean values, I explored the individual changes in motivation and additional L2 exposure for each participant in the study at different curricular levels. The individual changes in the participant's motivation (General Motivation, ideal L2 self, ought-to L2 self, and learning experience, respectively) and the changes in the students' additional L2 exposure in Phases 2, 3, 4, and 5 are also represented visually.

²⁴ Phase 1 data were collected one week before the start of the summer program and did not include the additional exposure log.

²⁵ For the purpose of simplicity, I did not investigate the hours of different aspects of additional L2 exposure for each curricular level as it was beyond the scope of RQ 1 and would also require adding 16 new tables.

	Mean SD SE		SE	95%	5 CI						
				Lower	Upper						
Phase 1 (n = 52)											
OSpan Composite	80.79	9.85	1.58	77.60	83.99						
BDS	21.60	9.24	1.43	18.71	24.48						
Ideal L2 self	6.28	1.95	0.27	5.73	6.82						
Ought-to L2 self	2.85	1.96	0.27	2.30	3.41						
Anti-ought-to L2 self	4.81	1.99	0.28	4.26	5.36						
Learning experience	7.62	1.38	0.22	7.19	8.06						
General Motivation	7.81	1.17	0.16	7.48	8.14						
	Phase 2	(n = 34)									
General Motivation	7.59	1.24	0.23	7.11	8.06						
Additional L2 exposure	16.95	9.90	1.75	13.38	20.52						
	Phase 3	(n = 33)									
General Motivation	7.59	1.10	0.20	7.18	8.00						
Additional L2 exposure	17.86	12.02	2.10	13.45	22.27						
	Phase 4	(n = 29)									
General Motivation	7.36	1.50	0.29	6.77	7.96						
Additional L2 exposure	17.09	11.12	2.10	12.77	21.40						
	Phase 5	(n = 24)									
Ideal L2-self	7.16	2.03	0.42	6.28	8.04						
Ought-to L2 self	3.59	2.34	0.50	2.55	4.63						
Anti-ought-to L2 self	4.61	2.16	0.45	3.67	5.54						
Learning experience	7.83	1.45	0.30	7.21	8.44						
General Motivation	7.80	1.55	0.32	7.13	8.47						
Additional L2 exposure	14.28	9.48	1.98	10.18	18.38						

Table 7. Descriptive statistics and 95% CI for the ID data (four curricular levels combined)

Notes.

Phase 1	= before	Week 1	in the summer	program;
---------	----------	--------	---------------	----------

Phase 2 = Week 2;

Phase 3 = Week 4;

Phase 4 = Week 6;

Phase 5 = Week 8 (last week of the summer program);

OSpan = operation span task (0-100 scale),

BDS = backward digit span task (0-50 scale);

L2MSS = L2 Motivational Self System (0-9 scale);

General Motivation = survey adapted from a subset of the Attitudes/ Motivation Test Battery (0-9 scale);

Learning experience (Phase 1) = Russian learning experience prior to the summer program;

Additional L2 exposure = total hours / week for all Russian learning activities outside formal classes.

Motivation as a dynamic ID: Trends at four curricular levels combined

The participants' motivation right before and during the summer program remained relatively stable, based on the results of the two motivation surveys for the four curricular levels combined (see Table 7). The participants were highly motivated L2 Russian learners from the very start of the program, as illustrated by the box and whisker plot in Figure 12 which represents the dynamics of their motivation measured by the General Motivation survey (data collected one week before the summer program and four times on a bi-weekly basis during the program). Even though the participants' motivation to study Russian slightly decreased from Phase 1 (M = 7.81, SD = 1.17, 95% CI [7.48, 8.14]) to Phase 4 (M = 7.36, SD = 1.50, 95% CI [6.77, 7.96]), it went back to the Phase 1 level by the time of Phase 5 (M = 7.80, SD = 1.55, 95% CI [7.13, 8.47]). The 95% CI for the five phases of the study measured by the General Motivation survey overlapped, suggesting that there were no statistically significant changes in the participants' motivation over the course of the summer program measured by the General Motivation test battery for the four curricular levels combined.



Figure 12. Motivation as a dynamic ID (General Motivation), four curricular levels combined, Phases 1-5.

The results for the L2MSS survey distributed twice during the study (one week before and at the end of the summer program) also did not reveal any statistically significant changes within the participants' visions of their ideal L2 self, ought-to L2 self, and anti-ought-to L2 self for the four curricular levels combined. Among the three types of selves, the largest (but no significant according to the 95% CI) changes were observed for the scores for ideal L2 self and ought-to L2 self, which both increased from Phase 1 to Phase 5. The ideal L2 self scores went from M = 6.28, SD = 1.95, 95% CI [5.73, 6.82] in Phase 1 to M = 7.16, SD = 2.03, 95% CI [6.28, 8.04] in Phase 5; ought-to L2 self went from M = 2.85, SD = 1.96, 95% CI [2.30, 3.41] in Phase 1 to M = 3.59, SD = 2.34, 95% CI [2.55, 4.63]) in Phase 5.

The scores for anti-ought-to L2 self remained relatively stable in Phases 1 (M = 4.81, SD = 1.99, 95% CI [4.26, 5.36]) and 5 (M = 4.61, SD = 2.16, 95% CI [3.67, 5.54]). The scores for

the Russian learning experience were consistently high for both the participants' prior institutions, as shown by the results for Phase 1 collected before the start of the summer program (M = 7.62, SD = 1.38, 95% CI [7.19, 8.06]) and the results for Phase 5, which the participants self-reported during the last week of their studies in the summer program (M = 7.83, SD = 1.45,95% CI [7.21, 8.44]). In terms of the differences between the components of L2MSS, the scores for the ideal L2 self in Phases 1 and 5 were significantly higher than the scores for ought-to L2 self and anti-ought-to L2 self in Phases 1 and 5, based on the 95% CI for the mean scores.



L2MSS: FOUR CURRICULAR LEVELS COMBINED

Figure 13. Motivation as a dynamic ID (L2MSS), four curricular levels combined, Phases 1 and 5.

The participants also discussed their motivation to study Russian during the interviews (the Russian segments in the excerpts are bolded). Some of these changes were qualitative in nature (e.g., changes from the need to maintain proficiency in the beginning to the need for full cultural immersion at the end of the program). Other changes were quantitative (e.g., going from high motivation in the beginning to even higher levels of motivation at the end of the program). For example, one participant mentioned that their "motivation would turn from simply maintaining a working proficiency to now I want to like see the world with the Russian eyes" [Excerpt 1, Participant M, Level 4, General Motivation/Ideal L2 Self]. Another participant who had also been exposed to Russian prior to the summer program mentioned that their "motivation to get back into Russian was really high in the beginning and then throughout the course it got even higher because of how challenging the language is" [Excerpt 2, Participant V, Level 1, General Motivation]. These findings for Participants M and V were also consistent with the quantitative data from the two motivation surveys (L2MSS and General Motivation) for the four curricular levels combined. These results also showed that the participants were generally highly motivated learners when joining the program and that their motivation to study Russian remained at the high levels by the end of the summer program (see Table 7 and Figures 11 and 12).

After completing the summer program, becoming fluent in Russian became a more tangible goal for several interviewees. This finding was associated with the increased clarity of the ideal L2, remaining consistent with the quantitative data from the L2MSS survey distributed during Phases 1 and 5. One of the participants stated that, "I really wanted to speak Russian [in the beginning of the program]... and I think [the summer program] made this is a tangible goal and showed me the journey to become fluent" [Excerpt 3, Participant C, Level 2, Ideal L2 Self]. Most of the participants stated that their parents, friends, or professors at their institutions

generally supported their decision to study Russian, but did not push them to pursue or not to pursue their Russian studies, even though some of them recommended "something more practical like Spanish or French" [Excerpt 4, Participant L, Level 4, Anti-Ought-To L2 Self] or even said that learning Russian "is a weird thing" [Excerpt 5, Participant D, Level 1, Anti-Ought-To L2 Self] or that Russian is "a little bit less important now than it would have been a few years ago" [Excerpt 6, Participant C, Level 2, Anti-Ought-To L2 Self]. The general support of the students' decision to study Russian was also reflected in the quantitative data from the L2MSS survey and the ought-to-L2 self and anti-ought-to L2 self, which remained relatively low over on Phases 1 and 5 and significantly lower than the ideal L2 self for the four proficiency levels combined (see Table 7). It is also important to note that even though there were no statistically significant changes in the participants' ought-to L2 self, the score for this component of L2MSS increased over the course of the summer program.

Even though the interviewees generally regarded their learning experience in the summer program (another component of L2MSS) as positive, which is consistent with the quantitative survey data, some of the students mentioned that having the program conducted entirely online had its limitations. One of the students mentioned that having the program "completely online was definitely a difference ... when I studied Spanish in middle and high school that was all in person, so there were more opportunities to actually practice speaking the language with your peers and to get that aspect of learning the language" [Excerpt 7, Participant V, Level 1, Learning Experience/Additional L2 Exposure]. Most of the interviewees also mentioned that the program was challenging because of the amount of work in and outside their classes. One of the students said that the program "was pretty brutal... three hours of classes several hours of homework and more of studying every day" and that they had "a stack of flashcards like eight inches tall" [Excerpt 8, Participant D, Level 1, Learning Experience/Additional L2 Exposure]. The students also mentioned that their studies in the summer program were often vastly different from the studies at their previous/current institutions, as the summer program focused on speaking (which most of them viewed positively) and for several interviewees, their institutions focused more on grammar.

Phase	Mean	SD	SE	95% CI		Mean	SD	SE	95% CI		Mean	SD	SE	95%	6 CI
				Lower	Upper				Lower	Upper				Lower	Upper
Phase 1:	Curricular Level 1 $(n = 14)$					Curricular Level 2 ($n = 14$)					Curricular Level 3 ($n = 15$)				15)
OSpan Composite	81.27	12.47	3.76	72.90	89.65	79.83	9.83	2.84	73.59	86.08	78.30	8.58	2.71	72.16	84.44
BDS	21.25	11.41	3.29	14.00	28.50	22.08	8.16	2.36	16.90	27.27	21.00	9.07	2.62	15.24	26.76

 Table 8. Descriptive statistics and 95% CI for the ID data (WM, Curricular Levels 1-4)

Table 9. Descriptive statistics and 95% CI for the ID data (motivation, Curricular Levels 1-4)

Phase	Mean	SD	SE	95%	6 CI	Mean	SD	SE	95%	5 CI	Mean	SD	SE	95%	6 CI
				Lower	Upper				Lower	Upper				Lower	Upper
Phase 1:	Curricular Level 1 $(n = 14)$					С	Curricular Level 2 $(n = 14)$				Curricular Level 3 $(n = 15)$				
Ideal L2 self	5.94	2.13	0.57	4.72	7.17	6.70	2.03	0.54	5.53	7.87	5.81	2.05	0.53	4.68	6.95
Ought-to L2 self	2.19	1.80	0.48	1.15	3.22	2.50	2.10	0.56	1.29	3.71	3.07	1.82	0.49	2.02	4.12
Anti-ought-to L2 self	4.89	2.24	0.60	3.59	6.18	4.71	2.08	0.56	3.51	5.91	4.65	1.79	0.46	3.66	5.64
Learning experience	6.94	2.24	1.00	4.16	9.72	7.50	1.19	0.33	6.79	8.22	7.75	1.53	0.39	6.90	8.59
General Motivation	7.85	0.98	0.27	7.25	8.44	7.71	1.21	0.32	7.02	8.41	7.69	1.39	0.36	6.93	8.46
Phase 2:	C	urricul	ar Leve	11(n = 1)	10)	Curricular Level 2 ($n = 10$)					Curricular Level 3 ($n = 10$)				
General Motivation	7.66	1.37	0.43	6.68	8.64	7.03	1.16	0.44	5.96	8.10	7.92	1.08	0.34	7.14	8.70
Phase 3:	C	Curricu	lar Lev	el 1 ($n =$	9)	Curricular Level 2 ($n = 11$)					Curricular Level 3 ($n = 10$)				
General Motivation	7.55	1.50	0.53	6.29	8.81	7.58	0.94	0.30	6.90	8.26	7.88	0.85	0.27	7.27	8.49
Phase 4:	C	Curricu	lar Lev	el 1 ($n =$	8)	Curricular Level 2 ($n = 11$)					Curricular Level 3 ($n = 7$)				
General Motivation	6.71	2.56	0.97	4.35	9.08	7.67	0.99	0.30	7.01	8.34	7.66	0.87	0.33	6.85	8.46
Phase 5:	C	Curricu	lar Lev	el 1 ($n =$	7)	Curricular Level 2 ($n = 8$)					Curricular Level 3 ($n = 7$)				
Ideal L2 self	6.67	3.36	1.37	3.15	10.19	6.73	1.75	0.62	5.26	8.19	8.17	0.64	0.24	7.58	8.76
Ought-to L2 self	2.27	1.73	0.71	0.45	4.08	3.51	2.61	0.99	1.10	5.93	4.17	2.34	0.88	2.01	6.33
Anti-ought-to L2 self	4.40	2.36	0.96	1.92	6.88	4.65	2.90	1.02	2.23	7.07	4.83	1.41	0.53	3.52	6.13
Learning experience	7.20	2.11	0.80	5.25	9.15	8.25	0.96	0.34	7.45	9.05	7.74	1.26	0.48	6.58	8.91
General Motivation	6.90	2.71	1.11	4.05	9.75	8.15	0.83	0.29	7.45	8.85	8.23	0.53	0.20	7.73	8.72

Phase	Mean	SD	SE	95%	% CI	Mean	Mean SD SE 95%		6 CI	Mean	SD	SE	95%	6 CI		
				Lower	Upper				Lower	Upper				Lower	Upper	
Phase 2:	Curricular Level 1 $(n = 10)$					C	Curricular Level 2 $(n = 10)$					Curricular Level 3 $(n = 10)$				
Add. L2 exposure	10.63	8.21	2.74	4.31	16.94	18.23	9.67	3.06	11.31	25.14	21.15	10.37	3.28	13.73	28.57	
Phase 3:	Curricular Level 1 $(n = 9)$					Curricular Level 2 ($n = 11$)					Curricular Level 3 ($n = 10$)					
Add. L2 exposure	12.38	11.57	4.09	2.71	22.04	18.18	10.61	3.20	11.06	25.31	21.05	13.87	4.62	10.39	31.71	
Phase 4:	Curricular Level 1 $(n = 8)$					Curricular Level 2 ($n = 11$)					Curricular Level 3 $(n = 7)$					
Add. L2 exposure	10.76	12.97	4.90	-1.24	22.75	18.50	7.57	2.28	13.41	23.59	19.23	13.21	4.99	7.01	31.44	
Phase 5:	Curricular Level 1 $(n = 7)$					Curricular Level 2 $(n = 8)$					Curricular Level 3 $(n = 7)$				7)	
Add. L2 exposure	13.65	12.31	5.02	0.74	26.56	15.38	10.43	3.69	6.65	24.10	14.66	8.22	3.11	7.06	22.26	

 Table 10. Descriptive statistics and 95% CI for the ID data (additional L2 exposure, Curricular Levels 1-4)

Motivation as a dynamic ID: Trends at different curricular levels

The dynamics of the students' motivation were also explored at the curricular levels separately (see Table 9 and Figure 14, 15, 16, 17, 18, and 19). These analyses showed one of the limitations of the present study – the discrepancies in the number of students who participated in different phases of the study. Even though 52 out of the 111 students enrolled in the summer program of 2021 initially participated in the study during Phase 1, the number of participants gradually decreased from Phase 2 (n = 34) to Phase 3 (n = 33), Phase 4 (n = 29), and Phase 5 (n = 24). Considering that the participants were enrolled in one of the four initial proficiency levels, the results for Phases 2-5 for separate curricular levels need to be interpreted with caution. The number of students from Curricular Level 4 was particularly low (n = 4, n = 3, n = 3, n = 2 in Phases 2, 3, 4, and 5 respectively). Statistically speaking, the low number of participants in Level 4 resulted in broad 95% CI for the mean scores that crossed zero in several cases and thus did not provide a precise estimate of the population means for the IDs. Due to these limitations, I decided not to conduct separate analyses of Curricular Level 4 for RQs 2 discussed in Chapters 7. This part of Chapter 6 also focuses only on Levels 1, 2, and 3.

Some of the patterns in the participants' motivation identified for the four curricular levels combined were also observed at the individual levels. The students at Curricular Levels 1-3 joined the program as highly motivated individuals, as indicated by the mean scores on the General Motivation survey on Phase 1 (M = 7.85, SD = 0.98, 95% CI [7.25, 8.44] for level 1; M = 7.71, SD = 1.21, 95% CI [7.02, 8.41] for Level 2; and M = 7.69, SD = 1.39, 95% CI [6.93, 8.46] for Level 3). The participants maintained approximately the same levels of motivation throughout Phases 1-5. The motivation as measured by the General Motivation survey for the participants who joined the program at Curricular Level 1 slightly decreased by the time of Phase

5 (M = 6.90, SD = 2.71, 95% CI [4.05, 9.75]). For Curricular Level 2 (M = 8.15, SD = 0.83, 95% CI [7.45, 8.85]) and Curricular Level 3 (M = 8.23, SD = 0.53, 95% CI [7.73, 8.72]), their motivation went up by the time of Phase 5 compared to the scores on Phase 1. Similarly to the results for the four curricular levels combined, the 95% CI for the five phases of the study measured by the General Motivation survey overlapped, suggesting that there were no statistically significant changes in the participants' motivation at Curricular Levels 1, 2, and 3 over the course of the summer program, as measured by this test battery.



Figure 14. Motivation as a dynamic ID (General Motivation), Curricular Level 1, Phases 1-5.



Figure 15. Motivation as a dynamic ID (General Motivation), Curricular Level 2, Phases 1-5.



Figure 16. Motivation as a dynamic ID (General Motivation), Curricular Level 3, Phases 1-5.

The data from the L2MSS survey distributed during Phases 1 and 5 were also analyzed for Curricular Levels 1, 2, and 3. Some of the L2MSS trends that were identified for the four curricular levels combined were also observed at the individual levels. As shown in Table 9 and Figures 17, 18, and 19, the participants' visions of ideal L2 self, ought-to L2 self, and anti-ought L2 self as measured by the L2MSS survey on Phase 1 at Curricular Levels 1, 2, and 3 were approximately at the same level when the students joined the program. As for the differences between the four components of L2MSS, even though the sample size was low at the individual curricular levels, the scores for ideal L2 self on Phase 1 for Levels 1, 2, and 3 were significantly higher than the scores for ought-to L2 self (M = 2.19, SD = 1.80, 95% CI [1.15, 3.22] for Level 1; M = 2.50, SD = 2.10, 95% CI [1.29, 3.71] for Level 2; and M = 3.07, SD = 1.82, 95% CI [2.02, 4.12] for Level 3), based on the 95% CI. The 95% CI also indicated that the Curricular Level 1 scores for anti-ought-to L2 self (M = 4.89, SD = 2.24, 95% CI [3.59, 6.18]) were significantly higher than their scores for ought-to L2 self.

By the time of Phase 5, the scores for ideal L2 self increased compared to Phase 1 for Curricular Levels 1 (M = 6.67, SD = 3.36, 95% CI [3.15, 10.19]) and 3 (M = 8.17, SD = 0.64, 95% CI [7.58, 8.76]) and remained almost the same for Curricular Level 2 (M = 6.73, SD = 1.75, 95% CI [5.26, 8.19]). The difference between the ideal L2 self for Curricular Level 3 on Phases 1 and 5 was statistically significant, based on the 95% CI that did not overlap (this finding is also illustrated in Figure 19). In terms of ought-to L2 self, the scores for this component of L2MSS slightly increased by the time of Phase 5 for Curricular Level 2 (M = 3.51, SD = 2.61, 95% CI [1.10, 5.93]) and Level 3 (M = 4.17, SD = 2.34, 95% CI [2.01, 6.33]), and remained at approximately the same for Curricular Level 1 (M = 2.27, SD = 1.73, 95% CI [0.45, 4.08]). The scores for anti-ought-to L2 self remained stable over the course of the summer program for

Curricular Levels 1, 2, and 3, as measured by the L2MSS survey on Phases 1 and 5. The participants' Russian learning experience for Phase 1 at their previous institutions and experience in the summer program on Phase 5 remained stable and high.



L2MSS: CURRICULAR LEVEL 1

Figure 17. Motivation as a dynamic ID (L2MSS), Curricular Level 1, Phases 1 and 5.



L2MSS: CURRICULAR LEVEL 2

Figure 18. Motivation as a dynamic ID (L2MSS), Curricular Level 2, Phases 1 and 5.



L2MSS: CURRICULAR LEVEL 3

Figure 19. Motivation as a dynamic ID (L2MSS), Curricular Level 3, Phases 1 and 5.

The qualitative interview data were collected from the students at different curricular levels. The general trends identified in the quantitative survey data collected during the summer were observed in the interviews that were recorded during the fall 2021 semester after the program. Remaining consistent with the quantitative data, several interviewees at different curricular levels stated that they joined the summer program as highly motivated Russian language learners and that their motivation to study the language remained stable and often increased by the end of the program. For example, one of the participants who was initially placed in Level 1 even though he had some prior Russian learning experience said that "it [motivation] has increased a little bit, although maybe not as much as others' simply because I had to study Russian before I went [to the summer program] as part of my program and I already knew I was pretty interested" [Excerpt 9, Participant C, Level 1, General Motivation]. Despite the differences in Russian language proficiency and prior experience learning the language, a similar pattern regarding the high initial motivation and even higher motivation by the end of the program was reported at Curricular Level 4. One of the interviewees from Level 4 mentioned that the 2021 program was their second program at that institution and said that "I had a pretty strong motivation coming in... I had a goal just to get as advanced as possible and I think over the course of the summer I just continued and continued pursuing that goal and afterwards I think my motivation has even increased more" [Excerpt 10, Participant L, Level 4, General Motivation].

The participants at different curricular levels also talked about the changes in their vision of ideal L2 self. One of the participants from Curricular Level 1 mentioned that "I do envision myself as someone who will speak Russian in the future and... I guess at the beginning of [the summer program] it was probably... I'll speak it for my job, but if I don't speak it for my job, I'll still be learning it and watching movies in Russian or something like that, just for fun" [Excerpt 11, Participant A, Level 1, Ideal L2 Self]. Based on the interview data from Curricular Levels 3 and 4, students at more advanced Russian proficiency levels obtained a clearer vision of someone who can speak Russian, remaining consistent with the quantitative data that includes the statistically significant changes in the scores for ideal L2 self. One of the interviewees from Level 4 mentioned that "I definitely envision myself as a person that will speak Russian in the future, and I hope to eventually use it for a career... I hope to use it every day for that, and I think through the course of [the summer program] that's just kind of confirmed" [Excerpt 12, Participant L, Level 4, Ideal L2 Self/Ought-To L2 Self]. The qualitative data analysis showed that students at different stages of their Russian L2 learning had similar visions of their ideal L2 self as illustrated by the quotations from Participants A and L: "I do envision myself as a person that will speak Russian in the future" (Level 1) and, "I definitely vision myself as a person that will speak Russian in the future" (Level 1).

Even though the participants' vision of the ideal L2 self had generally increased as indicated by the quantitative and qualitative data, one of the participants mentioned that "I've always been kind of realistic in the fact that my Russian would not get to the level that it would need to be to work in a professional context with... how I'm currently going to live my life" [Excerpt 13, Participant S, Level 3, Ideal L2 Self]. Some participants also reflected on the differences in Russian learners' motivation that they observed over the course of their studies. One of the interviewees from Curricular Level 3 mentioned that there are "some people who learn languages just to be proficient... they don't actually speak it, but I definitely like learning Russian in the beginning, and I really wanted to speak Russian" [Excerpt 14, Participant C, Level 3, General Motivation/Ideal L2 Self].

The participants also discussed how the summer program helped them re-evaluate their view of Russian as a subject, which is related to the concept of ought-to L2 self, since some of the participants joined the program because learning Russian would help them in their future careers. One of the summer program's graduates mentioned that "in terms of how it's [motivation] changed, I've really come to appreciate Russian as a language separate from how it can help me in my career path and I really think it's very interesting on top of obviously being very important" [Excerpt 15, Participant A, Level 1, Ought-To L2 Self]. It is important to note that some of the participants at lower proficiency levels also had their clear established goals and visions with regards to Russian when they started the summer program and maintained these goals and visions throughout the program and afterwards during the fall 2021 semester. This is an example related to ought-to L2 from a student who was initially placed in Level 1. This student considered pursuing a Ph.D. in Soviet history and thought that learning Russian would assist in achieving this goal. After talking about their professional goal, the participant stated that "I don't have a lot of intrinsic motivation to learn Russian" [Excerpt 16, Participant D, Level 1, Ought-To L2 Self]. The participant also added that "I don't enjoy learning languages as something just to do for fun... I find it really difficult and headache inducing" [Excerpt 17, Participant D, Level 1, General Motivation/Learning Experience]. This student's motivation to learn Russian was conceptually different compared to that of the other interviewees (entirely extrinsic motivation vs a combination of intrinsic and extrinsic factors). Despite these differences, the motivation of Participant D to learn Russian and their vision of someone who may not speak Russian in the future ("I probably will not be a person who speaks much Russian in the future because I am not willing to" [Excerpt 18, Participant D, Initial Curricular Level 1, General Motivation/Ideal L2 Self]) remained stable over the course of the summer program. A

similar stable but high motivation pattern was reported by some of the other interviewees. For example, Participant C joined the program as a highly motivated Russian language learner and maintained high levels of motivation throughout the program.

Motivation as a dynamic ID: Trends at the participant level

In addition to the mean values, I explored the individual changes in motivation and additional L2 exposure for each participant for Curricular Levels 1, 2, and 3. Figures 20, 21, 22, 23, and 24 represent the individual changes in motivation (General Motivation, ideal L2 self, ought-to L2 self, anti-ought-to L2 self, and learning experience, respectively). It is important to note that the number of participants who participated in different phases varied (see Table 7); there were also several instances when the students skipped some of the questions on the two motivation surveys. However, the graphs helped identify some important trends in the quantitative with regards to the individual participants. As it is seen in Figure 20, the individual trends in motivation as measured by the General Motivation survey at the participant level were more stable at Curricular Levels 2 and 3 compared to Level 1. Even though some of the participants from Curricular Level 1 maintained their high levels of motivation throughout the program, there were several participants whose motivation fluctuated and one participant whose motivation gradually decreased and reached a low level by the time of Phase 5, emphasizing the importance of investigating motivation at the individual participant level²⁶. Even though on average, the participants were considered highly motivated L2 Russian learners when they joined the program, the analysis at the individual participant level illustrated that their motivation ranged from moderate to high at Curricular Levels 2 and 3.

²⁶ This finding will be discussed in more detail in Chapter 8 in line with the principles of learner-centered pedagogy and learner-centered research.



GENERAL MOTIVATION: CURRICULAR LEVELS 1-3



Figure 20. Motivation as a dynamic ID (General Motivation), Curricular Levels 1-3, Phases 1-5.



IDEAL L2 SELF: CURRICULAR LEVELS 1-3

Figure 21. Motivation as a dynamic ID (ideal L2 self), Curricular Levels 1-3, Phases 1 and 5.



OUGHT-TO L2 SELF: CURRICULAR LEVELS 1-3

Figure 22. Motivation as a dynamic ID (ought-to L2 self), Curricular Levels 1-3, Phases 1 and 5.



ANTI-OUGHT-TO L2 SELF: CURRICULAR LEVELS 1-3

Figure 23. Motivation as a dynamic ID (anti-ought-to L2 self), Curricular Levels 1-3, Phases 1 and 5.



LEARNING EXPERIENCE: CURRICULAR LEVELS 2-3

Figure 24. Motivation as a dynamic ID (learning experience), Curricular Levels 2 and 3, Phases 1 and 5.

Note. Curricular Level 1 was not included in this figure because the beginner learners enrolled in the program had no or minimal prior L2 Russian learning experience.

The analyses of the L2MSS trends at the individual participant level also revealed some important results. There were several instances when the participants' ideal L2 self went from relatively high to low (Curricular Level 1) and from low to high (Curricular Levels 1-2). The scores for the ideal L2 self in Phase 1 ranged from low to high at the three curricular levels (see Figure 21). In Phase 5, the scores for Levels 1 and 2 ranged from low to high, whereas at Curricular Level 3, the participants' ideal L2 self was high overall. This finding illustrated the program helped increase the clarity of the ideal L2 self for all participants at Curricular Level 3, which also explained the statistically significant changes in the participants' scores for the ideal L2 self from Phase 1 to Phase 5 at Level 3. The participants' ought-to L2 self (see Figure 22) and anti-ought-to L2 self (see Figure 23) ranged from low to moderate or relatively high in Phases 1 and 5 at Curricular Levels 1, 2, and 3, further emphasizing the need to investigate the students' motivation at the individual participant level. In terms of the changes in the participant' ought-to L2 self and anti-ought-to L2 self, the results showed that these two types of selves remained relatively stable over the course of the program (even though for some of the participants their scores decreased, while for the others, they increased). In other words, the low scores for these two types of selves remained low, whereas the moderate or high scores remained moderate or high over the course of the program, which is consistent with the analyses of the quantitative and qualitative data discussed earlier in this chapter. While the participants' L2 Russian learning experience ranged from moderate to high at their previous institution (data collected in Phase 1), the participants generally positively viewed their experience in the summer program with the scores ranging from relatively high to very high (data collected in Phase 5, see Figure 24).

L2 exposure as a dynamic ID: Trends at four curricular levels combined

The additional L2 exposure data included the participants' responses in 13 categories total (see Figure 9). Upon the preliminary analysis and my consultation with the program's faculty, I decided to exclude the data from one of the categories from further analyses (Item 5 - Homework in the Russian program), as the hours spent on homework during the summer program were both quantitatively and qualitatively different from the other categories used in the additional L2 exposure log. As for the quantitative differences, homework was the most time-consuming activity compared to the other categories. On average, the participants spent 14 hours on homework from Monday through Friday during Weeks 2 and 4 (Phases 2 and 3) and 11 hours during Weeks 6 and 8 (Phases 4 and 5). The students often spent more time on homework in the summer program than on the other L2 activities combined. As for the qualitative differences, homework was also directly linked to the participants' formal classes. The goal of the additional L2 exposure log was to explore how students learned Russian outside their formal classes. These observations and future directions about the use of the additional L2 exposure log are discussed in detail in Chapter 8.

First, the additional L2 exposure data were analyzed at the four curricular levels combined. In line with the trends in the motivation data, the participants at the four levels (see Table 7 and Figure 13) were consistently engaged in Russian learning activities outside their formal classes Monday through Friday during Weeks 2, 4, 6, and 8 (corresponding to Phases 2, 3, 4, and 5). Even though the log data indicated that the participants' additional L2 Russian exposure in mean total hours/week slightly decreased from Phase 1 (M = 16.95, SD = 9.90, 95%CI [13.38, 20.52]) to Phase 5 (M = 14.28, SD = 9.48, 95% CI [10.18, 18.38]), the 95% CI for the four phases overlapped. This finding suggested that there were no statistically significant

131
changes in the amount of additional L2 exposure over the course of the summer program at the four curricular levels combined.



Figure 25. Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), four curricular levels combined, Phases 2-5.

For exploratory and pedagogical purposes, I also added visual representations of how much time the participants spent on the 12 different activities outside their formal classes in the summer program during Phases 2, 3, 4, and 5 (see Tables 11, 12, 13, and 14 and Figure 25). The two most popular activities in Phases 2, 3, and 4 that the participants at the four levels combined were engaged in were: listening to music in Russian and watching TV/movies/streaming services in Russian (on average, 3 hours on Phases 2, 3, and 4 in the two categories). In Phase 1, watching TV/movies/streaming services in Russian was the most popular activity (also, 3 hours on average), while the second most popular activity was conversations with peers in Russian (2

hours on average). In Phases 2, 3, and 4, listening to music and watching TV/movies/streaming services were followed by conversations with peers in Russian and conversations with instructors in Russian. On average, the participants spent 2 hours on these conversational activities during Phases 2, 3, 4, and 5. In Phase 5, however, language learning apps were the fourth most popular activity. Except for the three categories (listening to music, participation in thematic clubs, and conversations with tutors in Russian), the participants had well established learning routines and spent approximately the same amounts of time for each of the other nine learning activities outside their formal classes over the course of the summer program. In Phase 2, the participants listened to twice as much music in Russian as they did in Phase 5 (the 95% CI for the two phases only slightly overlapped). The participants' participation in thematic clubs also decreased over the course of the program (from approximately 2 hours in Phase 2 to 30 minutes in Phase 5 with the difference being statistically significant). A different pattern was observed for the conversations with the tutors (who were also faculty members of the summer program). The participants reported that they spent on average 2 hours on Phase 4 on the conversations with the tutors, as opposed to 1 hour with them on Phases 2, 3, and 5.

	Mean SD SE			95%	6 CI
				Lower	Upper
Listening to music in Russian	3.38	2.95	0.52	2.31	4.44
Watching TV/movies/streaming services in Russian	3.03	2.10	0.37	2.27	3.79
Conversations with peers in Russian	1.80	2.21	0.39	1.00	2.59
Conversations with instructors in Russian	1.78	3.45	0.61	0.54	3.02
Using Russian while participating in thematic clubs	1.59	1.98	0.35	0.88	2.31
Communication via social media/email in Russian	1.28	1.02	0.18	0.92	1.65
Language learning apps to study Russian	1.02	1.82	0.32	0.37	1.68
Conversations with tutors in Russian	0.78	1.09	0.19	0.39	1.17
Reading Russian literature	0.75	1.48	0.26	0.22	1.28
Using Russian while playing sports or video games	0.57	1.24	0.22	0.12	1.02
Reading Russian news/blogs online	0.55	0.91	0.16	0.22	0.88
Attending Russian guest lectures	0.42	0.69	0.12	0.17	0.67

Table 11. Additional L2 Russian exposure in mean hours in descending order, Monday through Friday, Phase 2 (n = 34) (four curricular levels combined)

Notes. Phase 1 data were collected one week before the start of the summer program and did not include the additional exposure log. Phase 2 = Week 2.

Table 12. Additional L2 Russian	exposure in mean	hours in desc	ending order,	, Monday t	hrough
Friday, Phase 3 $(n = 33)$ (four c	urricular levels con	mbined)			

	Mean	SD	SE	95%	o CI
				Lower	Upper
Watching TV/movies/streaming services in Russian	3.23	2.87	0.52	2.17	4.28
Listening to music in Russian	3.05	2.83	0.51	2.01	4.09
Conversations with instructors in Russian	2.24	4.02	0.72	0.77	3.72
Conversations with peers in Russian	1.62	2.22	0.40	0.80	2.43
Language learning apps to study Russian	1.29	1.95	0.35	0.57	2.01
Using Russian while participating in thematic clubs	1.29	1.78	0.32	0.64	1.94
Communication via social media/email in Russian	1.24	1.37	0.25	0.74	1.74
Using Russian while playing sports or video games	1.10	2.66	0.48	0.12	2.07
Conversations with tutors in Russian	1.07	1.31	0.24	0.59	1.55
Reading Russian literature	1.03	2.02	0.36	0.29	1.77
Reading Russian news/blogs online	0.41	0.71	0.13	0.15	0.67
Attending Russian guest lectures	0.29	0.53	0.09	0.10	0.48

Note. Phase 3 = Week 4.

	Mean	SD	SE	95%	6 CI
				Lower	Upper
Listening to music in Russian	2.70	2.61	0.49	1.69	3.71
Watching TV/movies/streaming services in Russian	2.61	1.90	0.36	1.87	3.34
Conversations with instructors in Russian	2.17	3.69	0.70	0.73	3.60
Conversations with peers in Russian	1.99	2.72	0.51	0.93	3.04
Conversations with tutors in Russian	1.95	5.69	1.07	-0.26	4.16
Communication via social media/email in Russian	1.23	1.40	0.26	0.69	1.78
Language learning apps to study Russian	1.18	1.83	0.35	0.47	1.89
Using Russian while participating in thematic clubs	1.04	1.53	0.29	0.44	1.63
Reading Russian news/blogs online	0.79	1.45	0.27	0.22	1.35
Using Russian while playing sports or video games	0.61	1.31	0.25	0.10	1.12
Reading Russian literature	0.52	0.96	0.18	0.15	0.89
Attending Russian guest lectures	0.32	0.77	0.15	0.02	0.62

Table 13. Additional L2 Russian exposure in mean hours in descending order, Monday through Friday, Phase 4 (n = 29) (four curricular levels combined)

Note. Phase 4 = Week 6.

Table 14. Additional L2 Russian exposure in mean hours in descending order, Monday through Friday, Phase 5 (n = 24) (four curricular levels combined)

	Mean	SD	SE	95%	% CI	
				Lower	Upper	
Watching TV/movies/streaming services in Russian	2.98	3.20	0.67	1.59	4.36	
Conversations with peers in Russian	2.15	2.82	0.59	0.93	3.37	
Conversations with instructors in Russian	2.01	3.59	0.75	0.46	3.56	
Language learning apps to study Russian	1.83	2.59	0.54	0.71	2.95	
Listening to music in Russian	1.67	1.90	0.40	0.85	2.50	
Communication via social media/email in Russian	0.81	0.93	0.19	0.41	1.22	
Conversations with tutors in Russian	0.73	0.86	0.18	0.36	1.10	
Reading Russian literature	0.63	1.26	0.26	0.08	1.18	
Reading Russian news/blogs online	0.57	0.84	0.18	0.20	0.93	
Using Russian while participating in thematic clubs	0.52	0.85	0.18	0.16	0.89	
Using Russian while playing sports or video games	0.21	0.51	0.11	-0.01	0.43	
Attending Russian guest lectures	0.17	0.83	0.17	-0.19	0.53	

Note. Phase 5 = Week 8 (last week of the summer program).



Figure 26. Additional L2 Russian exposure in total hours in 12 categories, four curricular levels combined, Phases 2-5.

Remaining consistent with the quantitative data from the additional L2 exposure log, the interviewees mentioned that they were consistently engaged in Russian learning activities outside their formal classes in the summer program (e.g., using language learning apps, watching movies, playing video games, listening to music, etc.). One of the participants commented on the importance of L2 exposure outside formal classes, especially since the 2021 summer program was conducted online: "I just tried to immerse myself as much as I could because I wasn't there in-person on campus" [Excerpt 19, Participant V, Level 1, Additional L2 Exposure/Learning Experience]. Another participant talked about watching Russian YouTube channels in their free time: "I watched a lot of Russian YouTube videos just to hear other people speak it and … there's this channel called "Easy Russian" and they just interview people on the streets, and I watched that a lot" [Excerpt 20, Participant A, Level 1, L2 Additional L2 Exposure]. For some of

the interviewees, additional sources of L2 exposure were not just online tools, but also their Russian speaking friends. One of the participants said, "I have some Russian friends that I do try to speak a little bit with and who are very gracious with me" [Excerpt 21, Participant N, Level 2, Additional L2 Exposure]. As for the changes in the participants' additional L2 exposure, Participant N mentioned that they had often used the language learning app Duolingo, but then switched to online flashcard tools and tried to watch movies in Russian over the course of their studies in the summer program. Several participants (including Participant N) mentioned that they watched Russian movies during both their formal classes and their free time.

L2 exposure as a dynamic ID: Trends at different curricular levels

The quantitative data from the additional L2 exposure log were also analyzed at the individual curricular levels. As it was pointed out earlier in this chapter, the number of participants in Curricular Level 4 was low (n = 4, n = 3, n = 3, and n = 2 on Phases 2, 3, 4, and 5, respectively). Therefore, these data were not analyzed or described further in Chapters 6, 7, and 8. This part of Chapter 6 also focuses on Curricular Levels 1, 2, and 3. With regards to the additional L2 exposure, the analyses also showed that the SD and 95% CI for this ID category were very broad at the individual curricular levels, particularly on Phases 4 and 5. This limitation of the sample size and/or the instrument made it difficult to calculate the precise estimate for the population (the 95% CI for Curricular Level 1 on Phase 1 also crossed zero). These findings illustrated that the amount of additional L2 exposure that the participants were engaged in over the course of the summer program varied, particularly at Curricular Level 1. In light of these observations, this part of Chapter 6 focuses on Phases 2 and 3 only.

The analysis of the additional L2 exposure data revealed that the participants' total number of hours that they spent on Russian learning activities outside their formal classes was

137

higher at Curricular Levels 2 and 3 as opposed to those at Level 1 on Phases 2 and 3. On average, the students enrolled in Curricular Level 1 spent approximately 11 hours on Phase 2 (M = 10.63, SD = 8.21, 95% CI [4.31, 16.94]) and 12 hours on Phase 3 (M = 12.38, SD = 11.57, 95% CI [2.71, 22.04]). The students at Curricular Level 2 spent approximately 18 hours on both Phase 2 (M = 18.23, SD = 9.67, 95% CI [11.31, 25.14]) and Phase 3 (M = 18.18, SD = 10.61, 95% CI [11.06, 25.31]); the students at Curricular Level 3 spent approximately 21 hours on Phase 2 (M = 21.15, SD = 10.37, 95% CI [13.73, 28.57]) and Phase 3 (M = 21.05, SD = 13.87, 95% CI [10.39, 31.71] on Phase 3). These findings illustrate that the higher the participants' initial curricular level in the summer program was, the more hours they spent learning Russian outside their formal classes. The participants at Curricular Levels 1, 2, and 3 were also consistently engaged in additional L2 Russian exposure, as shown by the results for Phases 2 and 3; this is consistent with the trend that was identified for the four curricular levels combined and discussed earlier in this chapter.



Figure 27. Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 1, Phases 2-5.



Figure 28. Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 2, Phases 2-5.



Figure 29. Additional L2 exposure as dynamic ID (in mean total hours / week, Monday through Friday), Curricular Level 3, Phases 2-5.

Based on the analysis of the qualitative data from the interviews, there were no major differences in types of additional L2 exposure that the students at different curricular levels were engaged in, even though the quantitative data showed that the participants at higher proficiency levels generally spent more time on Russian learning activities outside their formal classes in the summer program. The interviewees at all levels mentioned that they listened to music and watched movies in Russian, which is consistent with the quantitative data that showed that the two most popular activities that the participants at the four levels combined were listening to music in Russian (on average, 3 hours on Phases 2, 3 and 4 and 2 hours on Phase 5) and watching TV/movies/streaming services in Russian (on average, 3 hours on Phases 2, 3, 4, and 5). Several participants also mentioned podcasts. One of the interviewees who was initially enrolled in Level 1 said that "there's a podcast I'm listening to called "Russian Progress" with a

guy named Artem Nazarov I think... he has a whole bunch of episodes where he just goes through about his life and about how he learned to be a polyglot" [Excerpt 22, Participant V, Level 1, Additional L2 Exposure]. Participant V also explained that they were able to understand Nazarov's podcast better by the end of the program. At Curricular Level 2, one of the interviewees also talked about watching movies in Russian: "I think that one of the most useful activities that I've done was the Russian movies... other than the ones that the teachers were requiring us to watch... more contemporary movies I found on Netflix" [Excerpt 23, Participant R, Level 2, Additional L2 Exposure].

One of the interviewees that was initially placed in Curricular Level 3 mentioned the video game series "Animal Crossing" that was also a source of additional L2 Russian exposure: "I intentionally bought the "Animal Crossing" game [set in Russian] and the console for it because I know that most of my Russian expression stuff that I do outside of academic setting is gossip... I just gossip and when you're gossiping, you're not talking about couches or ovens or holes in the ground or other kind of boring little physical items that you would expect a first-year student to know" [Excerpt 24, Participant S, Level 3, Additional L2 Exposure]. One of the interviewees from Curricular Level 4 also talked about video games: "I try as much as possible to play video games in Russian, and I finished "Doom" in Russian and ? that was fun" [Excerpt 25, Participant M, Level 4, Additional L2 Exposure]. Participant M also mentioned language learning apps: "I'm kind of over language apps... it's too predictable and if you know how games work you can just game it and then it doesn't become as much of a learning exercise as it ought to be" [Excerpt 26, Participant M, Level 4, Additional L2 Exposure].

141

L2 exposure as a dynamic ID: Trends at the participant level

In addition to the mean values, I explored the individual changes in the participants' additional L2 exposure for each student in Curricular Levels 1, 2, and 3. Figure 30 represents the changes in the participants' additional L2 exposure in Phase 2, 3, 4, and 5. The analyses showed some important trends in the quantitative data at the individual participant level. The amount of additional L2 exposure in Phases 2, 3, 4, and 5 ranged from 5 to 40 hours at the three levels. Even though the data underwent several stages of cleaning and outlier analysis, some of the higher results seemed implausible, even though the mean numbers were reasonable as it was confirmed by the program's director and faculty. Some of the students may have overestimated the number of hours that they were engaged in learning Russian outside their formal classes in the summer program. This finding is considered a limitation of this dissertation research and discussed in Chapter 8. It is important to note that most of the results ranged from 5 to 20 hours of additional exposure per week, which was reasonable considering the immersive aspect of the summer program. Even though there were instances when the number of hours of additional L2 exposure that the students were engaged in increased or decreased (particularly at Curricular Levels 1 and 3), the numbers for each participant remained relatively stable, which is consistent with the results discussed earlier in this chapter. Overall, the students were actively and consistently engaged in learning Russian outside their formal classes in the summer program.



ADDITIONAL L2 EXPOSURE: CURRICULAR LEVELS 1-3

Figure 30. Additional L2 exposure as a dynamic ID, Curricular Levels 1-3, Phases 2-4.

Summary of Chapter 6

Chapter 6 included a summary of the analyses and results for RQ 1 (A and B), focusing on the longitudinal changes in the participants' motivation assessed by the L2MSS and General Motivation surveys and their additional L2 exposure in 12 different categories assessed by the biweekly log focusing on different Russian learning activities outside formal classes in the summer program. The data were analyzed at the four curricular levels combined, separately at the different levels, and at the individual participant level. The analyses of the quantitative data showed that the participants' motivation right before and during the summer program remained relatively stable, based on the results of the L2MSS and General Motivation surveys for the four curricular levels combined. Overall, the participants were highly motivated Russian language learners from the beginning of the program, which was also supported by the qualitative data from the interviews that were recorded in the Fall of 2021 after the summer program. Some of the patterns in the participants' motivation identified for the four curricular levels combined were also found at the individual Curricular Levels 1, 2, and 3 (due to the low number of participants in Curricular Level 4, no further analyses were conducted for this level). The students at Curricular Levels 1 through 3 joined the program as highly motivated individuals, as indicated by the mean scores on the General Motivation survey on Phase 1.

The results for the L2MSS survey that was distributed one week before the summer program and during week 8 at end of the program did not reveal any statistically significant changes within the participants' visions of their ideal L2 self, ought-to L2 self, and anti-ought-to L2 self for the four curricular levels combined. However, the difference between the scores for ideal L2 self for Curricular Level 3 on Phases 1 and 5 was statistically significant, based on the 95% CI. This finding was also supported in the analyses of the quantitative data at the individual

144

participant level. The graph for the ideal L2 self illustrated the program helped increase the clarity of the ideal L2 self for all participants at Curricular Level 3. The scores for ought-to L2 self and anti-ought-to L2 self ranged from low to moderately high, emphasizing the importance of investigating motivation at the individual participant level. However, these two types of selves remained stable (low, moderate, or relatively high) and lower compared to the scores for ideal L2 self for Curricular Levels 1, 2, and 3. For the four curricular levels combined, the scores for the ideal L2 self in Phases 1 and 5 were significantly higher than the scores for ought-to L2 self and anti-ought-to L2 self in Phases 1 and 5, based on the 95% CI for the mean scores. The participants' Russian learning experience for Phase 1 at their previous institutions and experience in the summer program on Phase 5 remained stable and high. The interviewees also supported this finding and generally regarded their learning experience in the summer program as positive. However, some of the students mentioned that having the program conducted entirely online had its limitations, which will be discussed in Chapter 8.

Remaining consistent with the quantitative data, several interviewees at different curricular levels stated that they joined the summer program as highly motivated learners and that their motivation to study the language remained stable and even increased by the end of the program. The qualitative data analysis showed that students at different stages of their Russian L2 learning had similar visions of their ideal L2 self: "I do envision myself as someone who will speak Russian in the future" (Level 1) and "I definitely vision myself as a person that will speak Russian" (Level 4). It is important to note that the analyses of the quantitative data at the individual participant level showed several instances when the students' motivation decreased over the course of the program. Considering this dissertation's orientation toward learner-centered pedagogy and research, this finding will also be discussed in Chapter 8.

145

Overall, the two most popular activities that the participants at the four curricular levels combined were engaged in were listening to music in Russian (on average, 3 hours on Phases 2, 3 and 4 and 2 hours on Phase 5) and watching TV/movies/streaming services in Russian (on average, 3 hours on Phases 2, 3, 4, and 5). The analysis of the additional L2 exposure data showed that the participants' total number of hours that they spent learning Russian outside their formal classes in the summer program was higher at Curricular Levels 2 and 3 when compared to Level 1 on Phases 2 and 3 (the additional L2 exposure data were not analyzed further for Phases 4 and 5 due to the potential issues with the sample size and/or the instrument). Remaining consistent with the quantitative data from the additional L2 exposure log, the interviewees mentioned that they were actively and consistently engaged in Russian learning activities outside their formal classes in the summer program (e.g., using language learning apps, watching movies, playing video games, listening to music, etc.). Based on the analysis of the interview data, there were no major differences in types of additional L2 exposure that the students at different curricular levels were engaged in, even though the quantitative data showed that the participants at higher proficiency levels spent more time learning Russian outside their formal classes. The interviewees at all levels mentioned that they actively listened to music and watched movies in Russian over the course of the program, which is consistent with the quantitative data.

CHAPTER 7: THE ROLE OF IDs IN L2 RUSSIAN LEARNING

Analysis overview of RQ 2

Bivariate correlation and multiple linear regression analyses were conducted to address the second RQ regarding the effects of WM capacity, motivation, and additional L2 exposure on the learning gains over the course of the summer program for L2 Russian lexicogrammar, speaking, and writing²⁷. The learning gains were calculated by subtracting the entrance scores from the exit scores on the lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs (the scores are shown in Table 15). Prior to calculating the learning gains, I verified that the students at each curricular level had approximately the same scores on the four placement tests. With some minor variation at Levels 2 and 3, the students at each curricular level had similar scores on the placement tests. After verifying the scores on the four tests, the descriptive statistics and 95% CI for the learning gains were calculated. The learning gains varied, depending on the participants' initial placement in the program (see Table 16). Except for the similar gains on the MC quiz and the WPTs at Levels 2 and 3, the students who joined the Russian language program at a lower level tended to have larger learning gains than the learners who started the program at a higher level and progressed more slowly on the program's internal FIB and MC quizzes, as well as the OPIcs and WPTs on the ACTFL scale represented by an inverted pyramid (for more information, see Merrill et al., 2021). As shown in Table 16, the students at Curricular Level 1 gained 22.50, 40.23, 4.08, and 4.42 points on average on the FIB quiz, MC quiz, ACTFL OPIcs, and ACTFL WPTs, respectively, whereas the participants at Level 2 gained 20.43, 22.93, 1.54, and 1.00 points and 9.31, 23.54, 1.08, and 1.27 at Level 3.

²⁷ The analyses and results for RQ 1 described in Chapter 6 revealed one of the limitations of this dissertation research—the number of students from Curricular Level 4 was particularly low (n = 4, n = 3, n = 3, n = 2 in Phases 2, 3, and 4, respectively). Therefore, I did not conduct separate analyses for Curricular Level 4 to address RQ 2.

Therefore, I conducted correlation analyses separately for Curricular levels 1, 2, and 3 to address RQ 2. I also included the participants' entrance scores as the predictors in the regression models with the exit scores as the outcome variables for the three levels combined. The general trends in the qualitative data regarding the relationships between the participants' IDs and their learning gains were described, including whether these trends aligned with the quantitative data. The main results identified in the analysis of the qualitative data were supported by quotations from the participants at different curricular levels.

For pedagogical and exploratory purposes, the students' L2 Russian development as measured by the ACTFL proficiency scale (rather than the numeric scale used in the analyses²⁸) with the four main levels (Novice, Intermediate, Advanced, Superior) subdivided into three sublevels (Low, Mid, and High) was also investigated for the OPIcs and WPTs. The results are shown in Tables 17 and 18 for speaking and writing, respectively. These results once again showed that students at the same initial proficiency levels often achieved different results, which makes the topic of IDs an important area of investigation in the field of ISLA and was a primary motivation for this dissertation research. For example, three students who joined the Russian summer program as Intermediate Mid speakers of Russian by the end were still Intermediate Mid speakers, while seven students reached the Intermediate High proficiency in speaking, and three students became Advanced Low speakers of Russian, as measured by the entrance and exit ACTFL OPIcs.

²⁸ To learn more about the numeric coding procedure for the OPIcs and WPTs, see Chapter 3.

	Entrance							Exit		
Test	Mean	SD	SE	95%	6 CI	Mean	SD	SE	95%	6 CI
				Lower	Upper				Lower	Upper
				Curricu	lar Level	1 (n = 1)	4)			
FIB	0.75	1.36	0.39	-0.11	1.61	23.25	12.21	3.53	15.49	31.01
MC	6.62	13.36	3.70	-1.46	14.69	46.85	14.83	4.11	37.88	55.81
OPIc	0.23	0.60	0.17	-0.13	0.59	4.33	0.49	0.14	4.02	4.65
WPT	0.07	0.27	0.07	-0.08	0.23	4.50	0.80	0.23	3.99	5.01
Curricular Level 2 ($n = 14$)										
FIB	21.50	14.45	3.86	13.16	29.84	41.93	17.37	4.64	31.90	51.96
MC	46.57	14.46	3.87	38.22	54.92	69.50	17.24	4.61	59.55	79.45
OPIc	3.93	1.07	0.29	3.31	4.55	5.46	0.78	0.22	4.99	5.93
WPT	4.61	0.93	0.31	3.90	5.32	5.77	0.88	0.26	5.18	6.36
				Curricu	lar Level	3(n = 1)	5)			
FIB	48.00	14.32	3.70	40.07	55.93	54.92	13.93	3.86	46.51	63.34
MC	76.00	19.02	4.91	65.47	86.53	99.15	16.77	4.65	89.02	109.29
OPIc	4.87	0.83	0.22	4.40	5.33	5.92	0.64	0.18	5.54	6.31
WP	5.46	0.50	0.14	5.14	5.77	6.08	0.56	0.16	5.73	6.44

Table 15. Descriptive statistics and 95% CI for the entrance and exit testing scores

Notes.

FIB = lexicogrammar fill-in-the-blank quiz;

MC = lexicogrammar multiple-choice quiz;

OPIc = Oral Proficiency Interview Computer,

WPT = Writing Proficiency Test.

	escriptive su	uisiics ana	9570 CI J	or the tearn	ing guins					
Learning	Mean	SD	SE	95%	o CI					
gains				Lower	Upper					
	Curr	icular Leve	11 (n = 14)	4)						
FIB	22.50	12.87	3.71	14.33	30.67					
MC	40.23	22.51	6.24	26.63	53.83					
OPIc	4.08	0.67	0.19	3.66	4.51					
WPT	4.42	1.00	0.29	3.78	5.05					
Curricular Level 2 ($n = 14$)										
FIB	20.43	13.12	3.51	12.85	28.00					
MC	22.93	11.45	3.06	16.32	29.54					
OPIc	1.54	0.88	0.24	1.01	2.07					
WPT	1.00	0.56	0.19	0.57	1.43					
	Curr	icular Leve	13 (n = 13)	5)						
FIB	9.31	11.18	3.10	2.55	16.06					
MC	23.54	12.63	3.50	15.91	31.17					
OPIc	1.08	0.76	0.21	0.62	1.54					
WPT	1.27	1.33	0.40	0.38	2.17					

J									
Placement					Exit O	PIc			
OPIc	NL	NM	NH	IL	IM	IH	AL	AM	AH
AB	0	0	0	7 (70%)	3 (10%)	0	0	0	0
NL	0	0	0	1 (100%)		0	0	0	0
NM	0	0	0	0	3 (100%)	0	0	0	0
NH	0	0	0	0	2 (100%)	0	0	0	0
IL	0	0	0	0	5 (55%)	4 (45%)	0	0	0
IM	0	0	0	0	3 (23%)	7 (54%)	3 (23%)	0	0
IH	0	0	0	0	0	0	0	0	0
AL	0	0	0	0	0	0	3 (50%)	3 (50%)	0
AM	0	0	0	0	0	0	0	0	0
AH	0	0	0	0	0	0	0	0	0

Table 17. Students' L2 speaking development during the summer program (number of students and % of all students in each level)

Notes. The participants who did not complete the program's exit OPI were not included in this table. AB = Absolute Beginner, N = Novice, I = Intermediate, A = Advanced; L = Low, M = Mid, H = High.

of all studer	ıts in e	ach lev	rel)						
Placement					Exit WP	Т			
WPT	NL	NM	NH	IL	IM	IH	AL	AM	AH
AB	0	0	0	0	2 (66%)	1 (33%)	0	0	0
NL	0	0	1 (100%)	0	0	0	0	0	0
NM	0	0	0	0	0	0	0	0	0
NH	0	0	0	1 (100%)	0	0	0	0	0
IL	0	0	0	0	1 (50%)	1 (50%)	0	0	0
IM	0	0	0	0	1 (8%)	10 (77%)	2 (15%)	0	0
IH	0	0	0	0	0	3 (75%)	1 (25%)	0	0
AL	0	0	0	0	0	0	1 (50%)	1 (50%)	0
AM	0	0	0	0	0	0	0	2 (100%)	0
AH	0	0	0	0	0	0	0	0	0

Table 18. Students' L2 writing development during the summer program (number of students and %

Notes. The participants who did not complete the program's exit WPT were not included in this table. AB = Absolute Beginner, N = Novice, I = Intermediate, A = Advanced; L = Low, M = Mid, H = High.

The correlation matrix with the Pearson's r coefficients and p values for the r effect sizes are shown in Tables 20, 21, 22 respectively for WM, additional L2 exposure, and motivation at Curricular Levels 1, 2, and 3. These matrices represent the relationships between the 21 IDs in different categories (independent variables) and the four learning gains on the FIB quiz, the MC quiz, the OPIcs, and the WPTs (dependent variables). In with previous research, (e.g., Plonsky & Oswald, 2014), the Pearson's r coefficients were interpreted as follows: close to .25 were considered small, .40 medium, and .60 large. The 95% CI were calculated for the large and statistically significant correlations to estimate the effect sizes for the population. For pedagogical and exploratory purposes, I also explored the correlations between the types of learning gains at different levels (see Table 19), which helped interpret some of the findings. It is important to note that the large and statistically significant positive correlations between the four types of learning gains were found at Curricular Level 1 only. At Curricular Levels 2 and 3, the only large and statistically significant positive correlations were found between the gains on the FIB quiz and the ACTFL WPTs. Due to the low number of observations for learning experience on Phase 1 (part of the L2MSS survey) at Curricular Level 1 (some of the students from Level 1 had no prior Russian learning experience), this variable was excluded from the correlation analyses at the separate curricular levels. The assumption of the absence of outliers was met for the correlations (outliers beyond three standard deviations from the mean were excluded prior to the main analyses; see Chapter 5). The level of measurement assumption of the Pearson's correlations was met (each variable used in the correlation analyses was continuous). The related pairs assumption was also checked (the pairwise deletion method was used throughout the analyses).

To further explore the role of IDs in L2 Russian learning, I conducted multiple linear regression analyses. In order to minimize the number of predictors used in the models due to the low sample size, the regression models for the WM and L2MSS included composite scores and a maximum of two predictors (an entrance FIB/MC/OPIc/WPT score and a composite ID score). The WM composite scores (including the data from Phase 1) were calculated from the OSpan composite and BDS scores (if a participant did not complete one of the WM tests and/or one of

151

the scores was excluded, this participant was excluded from the regression analyses). The L2MSS Composite scores were calculated for Phases 1 and 5 similarly to the WM composite scores. The effects of the scores on the General Motivation survey were explored in regression models for Phases 1, 2, 3, 4, and 5. The effects of the additional L2 exposure were explored for Phases 2, 3, 4, and 5. The multiple linear regressions analyses were conducted for the four exit scores (FIB quiz, MC quiz, ACTFL OPIc, and ACTFL WPT). The statistically and marginally significant predictors were also shown in scatter plots including the 95% CI. No evidence of multicollinearity was found in the regression models based on the variance inflation factors (VIF), which were included in the regression tables. The residuals in the regression models were normally distributed based on the Shapiro-Wilk tests. The assumption of homoscedasticity for the three models was checked by exploring the scatter plots with the standardized residuals and predicted values. The scatter plots showed random displacement, indicating that the homoscedasticity assumption was met.

Gains	1	FIB Gains	MC	OPIc Gains	WPT Gains						
			Gains								
		Curricular L	evel 1 (n =	= 14)							
FIB Gains	r		.61 *	.67*	.75**						
	р		.04	.02	.01						
MC Gains	r			.63*	.81 **						
	р			.03	.00						
OPIc Gains	r				.63*						
	р				.04						
WPT Gains	r										
	р										
Curricular Level 2 $(n = 14)$											
FIB Gains	r	1.00	.27	.12	.63						
	Р		.36	.70	.07						
MC Gains	r	.27	1.00	19	.17						
	р	.36		.54	.66						
OPIc Gains	r	.12	19	1.00	.45						
	р	.70	.54		.27						
WPT Gains	r	.63	.17	.45	1.00						
	р	.07	.66	.27							
		Curricular L	evel 3 (n =	= 15)							
FIB Gains	r	1.00	.47	08	.71*						
	р		.10	.79	.01						
MC Gains	r	.47	1.00	.20	.01						
	р	.10		.52	.98						
OPIc Gains	r	08	.20	1.00	.13						
	р	.79	.52		.70						
WPT Gains	r	.71*	.01	.13	1.00						
	р	.01	.98	.70							

Table 19. Relationships between the learning gains on the four tests

Notes. Statistically significant correlations at the .05 level are flagged with *. If a p-value is less than 0.01, it is flagged with **. Large correlations are bolded.

14010 20													
	Curricular Level 1 ($n = 14$)					Curricular Level 2 ($n = 14$)				Curricular Level 3 $(n = 15)$			
ID		FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT
OSpan	r	.32	.49	.22	.31	.05	.53	37	13	.57	26	09	.94**
	р	.34	.12	.54	.39	.89	.08	.27	.75	.09	.47	.80	<.001
BDS	r	.41	.48	.08	.43	.10	.15	28	25	.49	17	.03	.62
	р	.18	.11	.82	.19	.75	.64	.40	.52	.11	.60	.94	.06

Table 20. Relationships between WM and the learning gains on the four tests

Table 21. Relationships between additional L2 exposure and the learning gains on the four tests

	Curricular Level 1 (n = 14)					Cur	Curricular Level 3 $(n = 15)$						
Phase/ID		FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT
2 L2 exp.	r	.37	23	16	.17	.04	.05	.23	.37	.14	.11	05	17
	р	.33	.54	.67	.69	.92	.90	.55	.37	.70	.77	.90	.69
3 L2 exp.	r	.04	11	.06	.38	.32	12	.23	.29	08	.19	10	55
	р	.92	.80	.90	.40	.34	.73	.51	.45	.83	.63	.80	.20
4 L2 exp.	r	.17	09	38	.36	07	28	.45	.12	.63	.16	36	13
	р	.72	.85	.46	.49	.84	.40	.19	.77	.13	.73	.42	.81
5 L2 exp.	r	.35	.13	23	.54	.55	16	.38	.08	.55	.35	.41	.36
	р	.50	.81	.66	.34	.16	.70	.40	.86	.20	.44	.36	.48

		Curr	icular I	Level 1 (1	n = 14)	Curr	icular L	evel 2 (n	= 14)	Curricular Level 3 $(n = 15)$			= 15)
Phase/ID		FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT	FIB	MC	OPIc	WPT
1 Ideal	r	27	23	48	.05	.24	27	.32	.35	.13	04	08	.32
	р	.38	.45	.11	.88	.42	.35	.29	.36	.66	.90	.80	.34
Ought-to	r	.15	05	.38	.01	03	12	.37	.62	75**	.04	.45	69 *
-	р	.62	.87	.22	.97	.91	.69	.21	.07	.00	.91	.14	.03
Anti-ought-to	r	36	03	33	.01	17	.17	.16	48	.24	.26	.25	.02
	р	.22	.91	.30	.97	.56	.57	.61	.20	.42	.39	.42	.95
Learning exp.	r	51	.27	94	.34	16	06	.04	43	02	60 *	13	.25
	р	.38	.67	.06	.57	.61	.84	.89	.25	.95	.03	.66	.45
General Motivation	r	40	23	7 2**	27	.26	01	.66 *	11	26	27	.22	.13
	р	.18	.45	.01	.40	.37	.98	.01	.77	.40	.37	.47	.71
2 General Motivation	r	.11	37	17	17	48	.04	.29	53	44	.11	.55	.12
	р	.76	.30	.64	.66	.28	.94	.52	.36	.21	.75	.10	.78
3 General Motivation	r	.36	.46	.26	.73	.45	.01	.72*	.08	04	23	.02	.25
	р	.38	.25	.58	.07	.20	.98	.03	.85	.91	.51	.95	.55
4 General Motivation	r	.58	.37	.12	.77	.02	18	.43	23	10	38	.05	.30
	р	.18	.41	.82	.07	.96	.60	.22	.58	.84	.40	.92	.56
5 Ideal	r	.51	.59	.42	.79	.55	17	.53	02	.06	69	95**	.05
	р	.30	.22	.41	.11	.16	.69	.22	.97	.89	.09	.00	.93
Ought-to	r	.62	.62	.84*	.37	.40	19	.43	.16	91 **	38	37	73
	р	.19	.19	.04	.54	.38	.69	.39	.77	.01	.40	.41	.10
Anti-ought-to	r	01	15	02	.02	.14	37	.56	25	.18	38	38	.52
	р	.98	.77	.97	.97	.75	.37	.19	.58	.71	.40	.40	.30
Learning exp.	r	.47	.46	.23	.70	59	22	04	62	.55	.13	54	.33
	р	.29	.29	.62	.13	.13	.60	.93	.13	.20	.78	.21	.53
General Motivation	r	.78	.73	.46	.88	03	49	.55	28	.76*	14	11	.70
	р	.07	.10	.36	.05	.95	.22	.20	.54	.05	.76	.82	.12

Table 22. Relationships between motivation and the learning gains on the four tests

The role of WM in L2 Russian learning: Correlation analysis and interview insights

Bivariate correlation analyses revealed three large correlations between the scores on the WM tests and the learning gains at Curricular Level 3 (see Table 20). The strong statistically significant positive correlation was found between the OSpan Composite score and the learning gains in L2 Russian writing, as measured by the ACTFL WPTs (r = .94, p < .001, 95% CI [.84, .99]). The gains on the WPT also correlated with the scores on the BDS task (r = .62, p = .06, 95% CI [.38, .92]). The 95% CI for these effect sizes ranging from medium to large supported the facilitative role of WM, as measured by the two WM tests in the learning gains in L2 Russian writing at Curricular Level 3. Another strong correlation was found between the OSpan Composite score and the gains on the lexicogrammar FIB quiz (r = .57, p = .09, 95% CI [-.49, .95]). However, the 95% CI for the effect size in this case were broad and crossed zero, suggesting that this correlation was not representative of the population. The other correlations between the WM test scores and the four learning gains ranged from small to medium and were not statistically significant. The scores on the OSpan task tended to be positively correlated with the gains on the MC quiz at Curricular Level 1 (r = .49) and Level 2 (r = .53). The scores on the BDS task tended to be positively correlated with the gains on the FIB quiz (r = .41), the MC quiz (r = .48) and the WPT (r = .43) at Curricular Level 1. Another medium-to-large positive correlation between the BDS scores and the scores on the FIB quiz (r = .49) was found at Curricular Level 3. Overall, the quantitative results revealed multiple instances when a higher WM was associated with higher L2 Russian learning gains, particularly at Curricular Levels 1 and 3. The effects of WM on learning gains were explored further using multiple linear regression analysis for Phase 1. The results will be discussed later in this chapter.

During the interviews conducted after the summer program, the participants also talked about the benefits of having a good memory in general and the importance of the ability to quickly memorize new Russian words and grammar rules, which is associated with WM, remaining consistent with the general trends identified in the quantitative data. One of the participants from Curricular Level 1 mentioned that the ability to quickly memorize and store new material was particularly important because of the immersive aspect of the summer program: "... the program was so intensive because it forced you to hold on to certain details in order to move on to the next phase that you were learning" [Excerpt 27, Participant V, Level 1, WM/Learning Experience]. Participant V had WM scores slightly above the mean; the participant's FIB and WPT gains were also slightly above the means for Level 1, whereas the gains on the MC quiz and OPIcs were lower than the means. The interviewees also talked about different memorization strategies that they used during their Russian studies. Several participants mentioned rewriting and repetition, which was particularly important for them due to the fact the summer program was conducted entirely online in 2021. One of the students said that "having online material was a little harder for me... I found that when I was rewriting my exercises, I would remember the vocabulary and the grammar rules much better" [Excerpt 28, Participant R, Level 2, WM/Additional L2 Exposure]. Another participant had a summer job while also attending the summer program online. The participant mentioned that it was easier to memorize new Russian words and grammar rules before going to sleep: "I was grateful that I could do Russian before going to sleep" [Excerpt 28, Participant R, Level 2, WM/Additional L2] Exposure]. Several participants emphasized the importance of flashcards and notes, which helped them not only to memorize new words quickly but also to remember the words in the long term. One student from Curricular Level 3 mentioned: "I found that the most successful

thing for me to do is a lot of note cards, and I have to repeat them... I hear them and then I repeat them sometimes... in the mornings when I go to the gym, the first 20 minutes, I review my notes" [Excerpt 29, Participant S, Level 3, WM/Additional L2 Exposure]. Overall, the interviewees acknowledged that memory was a variable that all of them considered throughout their L2 Russian studies. The analysis of qualitative data also showed that many interviewees tried to develop strategies or routines that would help them memorize new Russian words and grammar rules more efficiently. It is important that these findings may be related to both WM/short-term and long-term memory abilities; however, since the importance of memory in general and WM specifically, as well as memorization strategies were mentioned by most of the interviewees, this topic will be further discussed in Chapter 8.

The role of motivation in L2 Russian learning: Correlation analysis and interview insights

The analysis of the relationships between motivation assessed by the L2MSS and General Motivation surveys and the four learning gains revealed both positive and negative large correlations (see Table 22): 14 large effects were found at Curricular Level 1; 8 large effects at Curricular Level 2; and 10 large effect sizes at Curricular Level 3 (32 large effects total). To minimize the probability of Type I error (i.e., rejecting the null hypothesis when it is true), only statistically significant large correlations were further investigated (10 correlations total).). After I identified the large effect sizes and statistically significant correlations, I calculated the 95% CI for the Pearson's r coefficients, similarly to the analyses for the WM tests and additional L2 exposure. Three large statistically significant correlations were found in Phase 1 at Curricular Level 3. The participants' ought-to L2 self at Level 3 negatively correlated with the gains on the lexicogrammar FIB quiz (r = -.75, p < .001, 95% CI [-.93, -.31]) and ACTFL WPTs (r = -.69, p = .03, 95% CI [-.92, -.10]). In other words, the higher the pressure from external sources was (e.g.,

from peers, teachers, and family members), the lower the students' learning gains on the FIB quiz and the WPTs were at Curricular Level 3. The 95% CI for the correlation between the ought-to L2 self and the gains on the FIB quiz ranged from large to medium, suggesting that this effect size was representative of the population; however, the 95% CI for the second correlation between the ought-to L2 self and the WPT gains were broader and ranged from a large to a small effect and thus were not considered representative of the population. It is important to note that the students' gains on the FIB quiz and WPT were significantly correlated at Curricular Level 3, which explained the similar correlation was also found at Level 3: the students' learning experience at their prior institutions was negatively correlated with the gains on the lexicogrammar MC quiz (r = -.60, p = .03, 95% CI [-.86, -.07]). Due to the broad confidence intervals that ranged from large to small effects, this correlation was not considered representative of the population was not considered representative of the population.

In Phase 1, one negative correlation at Curricular Level 1 (r = -.72, p = .01, 95% CI [-.92, -.24) and one positive correlation at Curricular Level 2 (r = .66, p = .01, 95% CI [.17, .89]) were found between General Motivation on Phase 1 and the students' gains on the lexicogrammar MC quiz. A strong positive correlation between General Motivation and the gains on the MC quiz was also found in Phase 3 of the study at Curricular Level 2 (r = .72, p = .03, 95% CI [.11, .94]). Considering the broad confidence intervals (particularly for the latter two correlations), these results need to be interpreted with caution.

The other four large and statistically significant correlations between motivation and the learning gains were found in Phase 5. Two negative correlations that were considered representative of the population based on the 95% CI were found between the components of

L2MSS and the learning gains were found at Curricular Level 3. The ideal L2 self was negatively correlated with the speaking gains on the ACTFL OPIcs (r = -.95, p < .001, 95% CI [-.99, -.69]). Another negative correlation was found between the ought-to L2 self and the gains on the lexicogrammar FIB quiz (r = -.91, p = .01, 95% CI [-.99, -.48]). The only positive strong correlation at Curricular Level 3 was found between General Motivation and the gains on the FIB quiz (r = .76, p = .05, 95% CI [.02, .96]); however, due to the broad 95% CI, this correlation was not considered representative of the population In Phase 5, there was also a positive strong correlation between ought-to L2 self and the OPIc gains at Curricular Level 1 (r = .84, p = .04, 95% CI [.09, .98]). The 95% CI for the effect size ranged from small to large effects and thus need to be interpreted with caution. In summary, the correlation analysis showed that the relationships between motivation as assessed by the scores on the L2MSS and General Motivation surveys and the learning gains in the summer program were complex; they varied at different proficiency levels; and they ranged from strong negative correlations to strong positive correlations. The role of motivation in L2 Russian learning was further explored in multiple regression analyses discussed later in this chapter.

Despite several nuances (for more detail, see Chapter 6), the participants at all four curricular levels generally talked about motivation (both intrinsic and extrinsic aspects of it) as a major factor in their studies in the summer program and before that, if applicable. Some of the interviewees considered motivation to be one of their strengths that helped them learn Russian. For example, a student who was initially placed in Curricular Level 2 said that "[My] strength ... I think [it is] motivation ... [it] definitely helps me out a lot because it... it keeps me from being burned out from wanting to study Russian" [Excerpt 30, Participant C, Level 2, General Motivation]. One of the important findings from the quantitative analysis was that there were several instances of strong negative correlations between different aspects of motivation and the students' learning gains in the summer program. Even though the interviewees did not talk about this phenomenon explicitly, one of the students mentioned that "You can be motivated as much as you want; you can have a great working memory, but if you're not using the language, if you're not exposed to it, it's not going to help you as much" [Excerpt 31, Participant V, Level 1, General Motivation/WM/Additional L2 Exposure].

The role of additional L2 exposure in L2 Russian learning: Correlation analysis and interview insights

The analyses revealed five large correlations between additional L2 exposure and the learning gains during the summer program (see Table 21). One large negative correlation (r =.55) was found in Phase 1, but it was not statistically significant. The other large effects were positive and found in Phases 4 and 5. The first strong positive correlation was found between the additional L2 exposure in Phase 4 and the students' learning gains on the lexicogrammar FIB quiz at curricular level 3 (r = .63, p = .13, 95% CI [.07, .96]). Considering the broad confidence intervals, this effect size was not considered representative of the population. Several strong positive correlations were also found in Phase 5 for Curricular Levels 1, 2, and 3; however, the 95% CI in these three cases were broad and crossed zero and thus were also not representative of the population. At Curricular Level 1, a large positive correlation was found between additional L2 exposure and the gains in L2 Russian writing, as measured by the ACTFL WPTs (r = .54, p =.34). Additional L2 exposure also positively correlated with the gains on the lexicogrammar FIB quiz at Curricular Level 2 (r = .55, p = .16) and Level 3 (r = .55, p = .20). Most of the other correlations between additional L2 exposure and the four types of learning gains were small. Two medium correlations were found for the speaking gains, as measured by the ACTFL OPIcs

in Phase 4 (r = .45) for Curricular Level 2 and in Phase 5 (r = .41) for Curricular Level 5. Overall, the analyses of quantitative data showed several instances when additional L2 exposure, particularly at later phases of the study, was facilitative of the learning gains at different curricular levels. Three out of four large positive correlations were found between additional L2 exposure and the learning gains on the lexicogrammar FIB quiz; however, none of these effects were statistically significant and representative of the population. The role of additional L2 exposure was further explored using multiple linear regression analyses discussed later in Chapter 6.

The analysis of qualitative data from the interviews showed that the participants were consistently engaged in different sources of additional L2 exposure during their studies (for more detail, see Chapter 6). In line with the quantitative data, when in several instances additional L2 exposure played a facilitating role in lexicogrammar gains, several interviewees mentioned that watching movies and listening to music in Russian outside the summer program helped them learn new words and grammar. One of the participants mentioned that in addition to movies, the language learning app Duolingo was very helpful: "I think that watching the movies and using Duolingo a little bit was very helpful... that really helped me to get a little more grammar and a little more words that I would not have otherwise probably ever really come into learning" [Excerpt 32, Participant C, Level 1, Additional L2 Exposure]. Participant C also mentioned the importance of English subtitles when watching movies in Russian and talked more about Duolingo, which the participant found particularly helpful because the summer program was conducted online in 2021: "...it's very different to use Duolingo... but I thought it kind of reinforced what I was learning in class in terms of pronunciation... in terms of skills that maybe I wasn't developing as much in the formal classroom setting just because of the format" [Excerpt

33, Participant C, Level 1, Additional L2 Exposure]. The interviewees who joined the summer program at higher proficiency levels also talked about a few of the benefits of additional L2 Russian exposure such as guest lectures and movies, particularly when it comes to learning slang and other informal expressions. According to one of the interviewees who was initially placed in Curricular Level 4 (the Russian segments are bolded), "the guest lectures were good and... and I now have a good collection of Russian movies to watch on my own but the lectures... especially the expressive lexicon... **awesome**... that was that was... **wonderful**" [Excerpt 34, Participant M, Level 4, Additional L2 Exposure]. Overall, the analysis of qualitative data showed that the participants from different curricular levels were consistently engaged in Russian learning activities outside their formal classes and tended to find these activities beneficial for their studies.

The role of WM, motivation, and additional L2 exposure in L2 Russian learning: Regression analyses

The multiple linear regression analyses were conducted separately for Phase 1-5, during which different ID data were collected. In line with the previous analyses, pairwise deletion was used in the regression analyses. Considering that the models included Curricular Levels 1-3, the entrance scores were used as the predictors (these scores were statistically significant predictors of the exit scores in each model). The analyses for Phase 1 revealed that L2MSS composite and General Motivation did not predict the exit scores on the four tests (these variables were not statistically significant and did not add value to the models); however, WM composite (see Tables 23 and 24) was a statistically significant predictor of the FIB exit scores ($\beta = .25$, p = .02) and the MC exit scores ($\beta = .25$, p = .01). These results are also represented visually in a form of scatter plots including 95% CI in Figure 31. Remaining consistent with the results of the

163

correlation analyses, the scatter plot for the FIB exit scores showed that a higher WM capacity was associated with higher scores at Curricular Levels 1 and 3; the scatter plot for the MC exit scores showed that a higher WM capacity was associated with higher scores at Curricular Level 1.

predictors of the ThD exit scores (Curricular Devels 1-5)											
	Unsta	andardiz	ed Coeffi	cients	S	tandardiz	Collinearity				
					C	Coefficien	Statistics				
	В	Std.	d. 95.0% CI (B)			t	Sig.	VIF			
		Error	Lower	Upper							
(Constant)	-5.05	12.12	-29.81	19.71		-0.42	.68				
FIB Entrance	0.66	0.09	0.48	0.84	0.78	7.61	<.001	1.00			
WM Composite	0.57	0.23	0.09	1.04	0.25	2.44	.02	1.00			
2											

Table 23. Regression model (Phase 1): FIB entrance scores and WM Composite as predictors of the FIB exit scores (Curricular Levels 1-3)

 $n = 32, R^2 \text{ model} = .69 \text{ (Adjusted } R^2 \text{ model} = .66, p < .001)$

Table 24. Regression model (Phase 1): MC entrance scores and WM Composite as
predictors of the MC exit scores (Curricular Levels 1-3)

Unsta	andardize	ed Coeffi	cients	St	andardiz	Collinearity	
				C	oefficie	Statistics	
В	Std.	95.0%	CI (B)	β	t	Sig.	VIF
	Error	Lower	Upper				
1.03	14.83	-29.26	31.32		0.07	.95	
0.69	0.07	0.55	0.84	0.85	9.56	<.001	1.00
0.78	0.28	0.21	1.34	0.25	2.80	.01	1.00
	Unst: B 1.03 0.69 0.78	Unstandardize B Std. Error 1.03 14.83 0.69 0.07 0.78 0.28	Unstandardized Coeffie B Std. 95.0% Error Lower 1.03 14.83 -29.26 0.69 0.07 0.55 0.78 0.28 0.21	Unstandardized Coefficients B Std. 95.0% CI (B) Error Lower Upper 1.03 14.83 -29.26 31.32 0.69 0.07 0.55 0.84 0.78 0.28 0.21 1.34	Unstandardized Coefficients Stress of the second sec	Standardize Standardize B Std. 95.0% CI (B) β t Error Lower Upper 0.07 0.07 1.03 14.83 -29.26 31.32 0.07 0.69 0.07 0.55 0.84 0.85 9.56 0.78 0.28 0.21 1.34 0.25 2.80	Unstandardize Coefficients Standardize B Std. 95.0% Γ (B) β t Sig. B Std. 95.0% Γ (B) β t Sig. Inor Lower Upper 0.07 .95 0.69 0.07 0.55 0.84 0.85 9.56 .001 0.78 0.28 0.21 1.34 0.25 2.80 .01

 $n = 32, R^2 \text{ model} = .76 \text{ (Adjusted } R^2 \text{ model} = .75, p < .001)$





Figure 31. WM Composite (Phase 1) as a predictor of the exit scores on the FIB and MC quizzes, Curricular Levels 1-3, error bars = 95% CI.

The multiple linear regression analyses for Phases 2-4 including General Motivation and additional L2 exposure did not reveal any statistically significant predictors. General Motivation and additional L2 exposure did not add value to the regression models for these three phases (for simplicity, these models were not included in this manuscript). Considering that the number of participants in Phase 5 was the lowest compared to the other phases, these results need to be interpreted with additional caution. In Phase 5 (see Table 25), General Motivation was a marginally significant predictor of the MC exit scores ($\beta = .24$, p = .05). This result is also represented visually in a form of a scatter plot including 95% CI in Figure 32. Remaining consistent with the bivariate correlation analyses discussed earlier in this chapter, the scatter plots illustrated that the General Motivation was particularly important for Curricular Level 1 in terms of the exit MC scores.

as predictors of the MC exit scores (Curricular Levels 1-5)											
	Unsta	St	tandardi	Collinearity							
					C	Coefficie	Statistics				
	В	Std.	95.0%	CI (B)	β	t	Sig.	VIF			
		Error	Lower	Upper							
(Constant)	10.35	15.70	-22.64	43.34		0.66	.52				
MC Entrance	0.67	0.10	0.47	0.87	0.81	7.02	<.001	1.01			
General Motivation	4.04	1.95	-0.05	8.14	0.24	2.07	.05	1.01			
	1 - 2										

Table 25. Regression model (Phase 5): MC entrance scores and General Motivation (Phase 5) as predictors of the MC exit scores (Curricular Levels 1-3)

 $n = 20, R^2 \text{ model} = .76 \text{ (Adjusted } R^2 \text{ model} = .73, p < .001)$



Figure 32. General Motivation (Phase 5) as a predictor of the exit scores on the MC quiz, Curricular Levels 1-3, error bars = 95% CI.

Summary of Chapter 7

The quantitative results revealed multiple instances when a higher WM as measured by the OSpan and BDS tasks played a facilitative role in the L2 Russian learning gains over the course of the summer program, specifically at Curricular Levels 1 and 3. Statistically significant large correlations were found between the scores on the WM tests and the learning gains at Curricular Level 3. The first strong correlation was found between the OSpan Composite score and the learning gains in L2 Russian writing (r = .94, p < .001, 95% CI [.84, .99]). Another strong correlation was found between the writing gains and the scores on the BDS task (r = .62, p = .06, 95% CI [.38, .92]). The 95% CI for the effect sizes supported the facilitative role of WM
in the L2 writing gains at Curricular Level 3. A strong correlation was also found between the OSpan Composite score and the gains on the lexicogrammar FIB quiz (r = .57, p = .09, 95% CI [-.49, .95]), but the 95% CI for the effect size were broad and crossed zero, suggesting that this correlation was not representative of the population. The qualitative data also supported the general trend regarding the facilitative role of memory in L2 Russian learning. The interviewees acknowledged that memory was a variable they considered throughout their L2 Russian studies in the summer program and that many of them tried to develop strategies or routines that would help them memorize Russian vocabulary and grammar more efficiently by taking their own IDs into consideration. The results of the multiple linear regression analyses also supported the finding regarding the facilitative role of WM in L2 Russian learning of lexicogrammar; however, the role was fairly small especially when compared to the influence of the pretest scores. The composite score of WM was a statistically significant predictor of the FIB exit scores ($\beta = .25$, p = .02) and the MC exit scores ($\beta = .25$, p = .01).

The relationships between motivation and learning gains were less straightforward than the positive correlations between WM capacity, additional L2 exposure, and the learning gains discussed in this chapter. Considering that there were 32 large correlations between different components of L2MSS and General Motivation, I decided to narrow down this number and discussed only large and statistically significant correlations with 95% CI that did not cross 0 and thus were representative of the population. After narrowing down the results, I identified 10 large and statistically significant correlations between motivation and the four types of learning gains. Strong negative correlations in Phase 1 were found between ought-to L2 self and the gains on the FIB quiz (r = -.75, p < .001, 95% CI [-.93, -.31]) and the ACTFL WPTs (r = -.69, p = .03, 95% CI [-.92, -.10]) at Curricular Level 3. The students' prior Russian learning experience on Phase 1 was also negatively correlated with the gains on the MC quiz (r = -.60, p = .03, 95% CI [-.86, -.07]) at Curricular Level 3. Due to the broad confidence intervals ranging from large to small effects, the latter two correlations were not considered representative of the population. Another negative correlation was found in Phase 1 at Curricular Level 1 between General Motivation and the gains on the speaking OPIcs (r = -.72, p = .01, 95% CI [-.92, -.24). At Curricular Level 2, General Motivation was positively correlated with the gains on the speaking OPIcs in Phase 2 (r = .66, p = .01, 95% CI [.17, .89]) and Phase 3 (r = .72, p = .03, 95% CI [.11, .94]). In Phase 5, another series of strong negative correlations was found at Curricular Level 3: ideal L2 self was negatively correlated with the OPIc gains (r = -.95, p < .001, 95% CI [-.99, -.69]); ought-to L2 self was negatively correlated with the gains on the FIB quiz (r = -.91, p = .01, 95% CI [-.99, -.48]). At Curricular Level 3 (Phase 5), there was also a positive correlation found between General Motivation and the gains on the lexicogrammar FIB quiz (r = .76, p = .05, 95% CI [.02, .96]). Finally, a strong positive correlation between ought-to L2 self and the gains on the OPIc was found in Phase 5 at Curricular Level 1 (r = .84, p = .04, 95% CI [.09, .98]). In summary, the correlation analysis showed that the relationships between the students' motivation and the learning gains in the summer program were complex; the effect sizes varied at different curricular levels and ranged from strong negative correlations to strong positive correlations. The results of the quantitative analysis supported the importance of motivation as an ID in L2 Russian learning research, which was also consistent with the qualitative data from the interviews. Despite the rich insight regarding the role of motivation provided in the interviews, the qualitative data did not shed light upon the several instances of strong negative correlations between different aspects of motivation and the students' learning gains. The multiple regression analyses revealed only one instance when motivation was marginally significant predictor of the

scores on the exit test. General Motivation was a marginally significant predictor of the MC exit scores ($\beta = .24$, p = .05). Due to the low sample size (particularly in Phase 5), this result needs to be interpreted with caution.

The analyses of quantitative data from the bi-weekly logs showed several instances when additional L2 exposure played a facilitative role in the learning gains at different curricular levels, particularly later in the program in Phases 4 and 5. Four large positive correlations between additional L2 exposure and the learning gains during the summer program were found; however, none of these correlations were statistically significant and were not considered representative of the population. The additional L2 exposure in Phase 4 correlated with the students' learning gains on the lexicogrammar FIB quiz at Curricular Level 3 (r = .63, p = .13, 95% CI [.07, .96]). At Curricular Level 1, a large positive correlation was found between additional L2 exposure and the gains in L2 Russian writing (r = .54, p = .34); additional L2 exposure also positively correlated with the gains on the FIB quiz at Curricular Level 2 (r = .55, p = .16) and Level 3 (r = .55, p = .20). Most of the other correlations between additional L2 exposure and the four types of learning gains were small at different curricular levels and phases of the study. The finding regarding the facilitative (even though not significant) role of additional L2 exposure (particularly for lexicogrammar) was supported by the qualitative data from the interviews. The interviewees from all curricular levels were consistently engaged in L2 Russian learning outside their formal classes in the summer program. Moreover, the students tended to find these activities beneficial for their studies. Despite the trend toward a facilitative role of additional L2 exposure in L2 Russian learning identified in the quantitative and qualitative data, the analyses showed that this ID was not a significant correlate of the four types of gains. This finding was also supported in the multiple linear regression analyses for Phases 2, 3, 4, and 5.

Additional L2 exposure was not a statistically significant predictor of the exit scores and did not add value to the models.

CHAPTER 8: GENERAL DISCUSSION AND CONCLUSION

Summary of the findings

General discussion

This longitudinal mixed-methods study of IDs in instructed L2 Russian learning over the course of the Russian summer immersion program was guided by DST (e.g., de Bot et al., 2007; Ellis, 2007; Larsen-Freeman, 2014) and multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013). The study focused on the three important categories of L2 learners' IDs: WM (as measured by the OSpan and BDS tasks), motivation (assessed by the L2MSS and General Motivation surveys), and differences in L2 exposure (specifically L2 Russian exposure outside formal classes assessed by the bi-weekly log). Despite the propositions regarding the interconnectedness of IDs (e.g., Denhovska et al., 2016; Dörnyei & Ryan, 2015; Segalowitz & Trofimovich, 2012), previous research has often explored these L2 learners' IDs separately and considered them as "modular" (i.e., discrete, separate, isolated), stable, and context independent traits" (Serafini, 2017, p. 370). In this dissertation research, I viewed WM, motivation, L2 exposure, and L2 learning gains as a complex system of multi-layered constructs rather than single-dimensional IDs and separate language skills.

In line with DST, motivation and L2 exposure were considered as dynamic constructs, with the quantitative data collected over five phases during the summer program and the qualitative data collected during the post-program interviews. The inclusion of different ID categories, multiple phases, and different types of data emphasized the need to reconsider the cognitive/affective dichotomy (see Dörnyei, 2010); addressed the need for more longitudinal ISLA research of IDs with multiple data points (Boo et al., 2015; Papi & Hiver, 2020; Martinsen

et al., 2011; Ranta & Meckelborg, 2013; Serafini & Sanz, 2016) and more studies with LCTLs such as Russian in ecologically valid settings (Denhovska et al., 2016; Pastushenkov & McIntyre, 2020); as well as revealed various benefits of the mixed-methods approach. This study's methodology, a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018), has been adopted in previous ISLA research that also considered the role of learners' IDs (e.g., Fathi et al., 2019; Pastushenkov & McIntyre, 2020; Saito & Hanzawa, 2018). In this dissertation research, this methodological approach helped obtain a unique view of L2 Russian learners' IDs, helped to better understand how and why motivation and additional L2 exposure fluctuated over time, and further explored the role of IDs in L2 learning gains for lexicogrammar, speaking, and writing. This approach also aligned with the idea of learnercentered research, which became one of the guiding principles of this study. In line with previous research (e.g., Lee & Revesz, 2021; Sagarra, 2017; 2019; Serafini, 2017), WM was viewed as a more stable cognitive construct with the data collected once immediately before the summer program, which is considered a potential limitation and discussed later in this chapter along with the other limitations such as the study's sample size, particularly in later phases.

The important role of context (i.e., the setting where the study was conducted), one of the main principles of DST (see, e.g., Waninge et al., 2014), guided the entirety of the study, including its design and analyses. Collecting data at a large Russian summer immersion program with four curricular levels (First-, Second-, Third-, and Fourth-Year Russian are referred to as Curricular Levels 1-4) imposed multiple challenges (for more detail, see Chapters 3 and 4) and resulted in a number of limitations. However, it was also a unique opportunity, considering that this intensive immersion program yielded learning gains comparable to one year of university instruction (see Merrill et al., 2021; Rifkin, 2005). The students at Curricular Level 1 gained

22.50, 40.23, 4.08, and 4.42 points on average on the FIB quiz (0-100 scale), MC quiz (0-150), ACTFL OPIcs (0-9 scale), and ACTFL WPTs (0-9 scale), respectively, whereas the participants at Level 2 gained 20.43, 22.93, 1.54, and 1.00 points and 9.31, 23.54, 1.08, and 1.27 for the same respective quizzes and tests at Level 3. Another important finding was that the learning gains varied. For example, three students who joined the summer program as Intermediate Mid speakers on the ACTFL scale were by the end still Intermediate Mid speakers, while seven students reached the Intermediate High proficiency, and three other students became Advanced Low speakers of Russian. Even though the analyses of the quantitative and qualitative data did not help explain the differences in the ACTFL OPIcs scores, the students' WM capacity and their motivation at the end of the program helped shed some light on the differences in the lexicogrammar quizzes and the ACTFL WPTs. Overall, this study illustrated that IDs not only helped explain some of the discrepancies in learning gains, but also that the analysis of the dynamics or changes in IDs such as motivation can serve as a proxy of the effectiveness of the program's pedagogy. The finding regarding the stability of the students' motivation over the course of the summer program at the four levels combined, as well as the important changes at the curricular and individual participant levels, will be discussed in detail later in this chapter and will also be considered in the pedagogical implications section.

Overall, the results of this study indicated that L2 learners' motivation and L2 exposure were constructs showing both stable (referred to as attractor states in DST literature, e.g., Hiver, 2015; Papi & Hiver, 2020) and non-stable dynamic behaviors that were often difficult to explain or predict (referred to as erratic patterns in DST literature, e.g., Serafini, 2017). Further emphasizing the need to consider multidimensional perspectives of L2 learners' IDs (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013), different constituents of the three IDs (WM, motivation,

and L2 exposure) and learning gains were engaged in complex relationships that varied at different curricular levels, language skills, and phases of the study. This discussion includes a section on theoretical implications with regards to DST and the mixed-methods approach as future directions for ISLA studies of IDs, including research with LCTLs. The dissertation also includes a section on practical implications that discusses pedagogical aspects and implications for the development of background questionnaires and placement testing that take L2 learners' IDs into consideration.

Discussion of the dynamics of motivation

With RQ 1 (A) discussed in Chapter 6, I aimed to explore to what extent L2 Russian students' motivation changed over the course of the summer program. In this study, I explored the changes in the "traditional" components of Dörnyei's (2009a) L2MSS (ideal L2 self, oughtto L2 self, and learning experience), in addition to the newly established anti-ought-to L2 self (see, e.g., Thompson, 2017, 2021), also referred to as the "rebellious" self (Lanvers, 2016). The anti-ought-to L2 self was added to the L2MSS survey, as previous research has shown that this type of selves was particularly important for L2 learners of LCTLs, including Russian, in Anglophone contexts (Pastushenkov & McIntyre, 2020; Thompson, 2017). In line with the study by Serafini (2017), the L2MSS survey (20 items total) was distributed twice: immediately before the summer program (Phase 1) and at the end during Week 8 (Phase 5). These quantitative data were then processed into eight independent variables (four on Phase 1 and four on Phase 5), as well the two L2MSS composite scores that were calculated later for the multiple linear regression analyses. In line with Serafini (2017), a longitudinal ID study of instructed L2 learning of Spanish that was also guided by DST and attempted to reconsider the cognitive/affective dichotomy (see Dörnyei, 2010), I added a second motivational framework-

a survey with five items adapted from the Motivation subset (effort, desire, and attitudes toward learning) of the AMTB developed by Gardner and colleagues (e.g., Gardner et al., 2004). Using two motivational frameworks has been a common practice in L2 motivation research (Boo et al., 2015), including ISLA research (Csizér, 2017), as this methodological aspect helped obtain a more fine-grained view of the role of motivation in L2 development. As I learned from the dissertation pilot study (Pastushenkov, 2020), for some of the students it was difficult to find time to participate in research due to the immersive aspect of the Russian summer program (for more detail, see Chapter 4). When working on the study design, I needed to find a balance between the number of phases of the study and the participants' availability. This is why the quantitative ID data were collected on a bi-weekly basis (even though a shorter period of time between the phases of an empirical study guided by DST may have been beneficial, see, e.g., de Bot, 2015). Considering that it was important to assess motivation at multiple data points (remaining consistent with the principles of DST, e.g., Larsen-Freeman, 2014), I tried to create short surveys that would help better understand the motivation of L2 Russian learners, which has remained an underexplored area of ISLA research similarly to the other LCTLs, despite the clear need for more studies in this area (Boo et al., 2015; Pastushenkov & McIntyre, 2020; Thompson, 2017).

A primary goal of the analyses of the quantitative and qualitative data was to identify trends in the dynamics of the students' motivation at the four curricular levels combined, as well as separately at the different curricular and participants levels, going from general patterns to more individualized observations. Based on the data from the L2MSS and General Motivation surveys, the participants' motivation over the course of the summer program remained relatively stable and high at the four curricular levels combined. The participants were highly motivated L2

Russian learners from the beginning of the program, which was also supported by the interview data. Theoretically speaking, this stable motivational pattern was considered an attractor state in DST terms (e.g., Papi & Hiver, 2020) as motivation (at least at the four levels combined) remained relatively stable over the course of the summer program. Pedagogically speaking, this finding indicated that the summer program was able to prepare engaging and enriching classes and helped maintain high levels of students' motivation, which was particularly important and difficult due to the COVID-19 pandemic.

There were also cases when the participants' motivation went from moderate or relatively high to even higher at the individual curricular levels. These results were consistent with the results of the case study by Pastushenkov and McIntyre (2020) conducted with Katia and Yana, two graduates of a Russian summer immersion program. Despite the differences in the case study's participants' ultimate attainment and their long-term engagement in L2 Russian learning, the quantitative and qualitative data from the case study showed the summer program helped both Katia and Yana increase the clarity of their ideal L2 selves. In the case of Katia, this strong ideal L2 self transformed into a long-term commitment to learning Russian. The increased clarity of the students' ideal L2 self (particularly at Level 3) is considered beneficial, as L2 motivation in general has been viewed as a vehicle for L2 learning (Sato & Loewen, 2019a) and has been associated with the effort that L2 learners are willing to invest in their studies (e.g., Csizér, 2017, Dörnyei, 2009a). Ideal L2 self has often been considered related to intrinsic motivation (i.e., "doing an activity for its inherent pleasure and enjoyment rather than external consequences," Lai, 2013, p. 94) and viewed as a constituent of L2 motivation that has been associated with long-term commitment to L2 learning (Pastushenkov & McIntyre, 2020), as well as learning effort and achievement (Dörnyei & Chan, 2013). If L2 proficiency is a part of a learner's vision

of the ideal L2 self, this individual will be motivated to acquire the L2 because of the desire to reduce the discrepancy between the current and future self (Ushioda & Dörnyei, 2012). A strong vision of the ideal self implies that L2 Russian learners can visualize as someone who can speak the language (Thompson, 2017), further emphasizing the benefits of immersive learning environments (see, e.g., Isabelli-García & Lacorte, 2016; Merrill et al., 2021) such as the eightweek Russian immersion program.

It is important to note that the items on the L2MSS survey focusing on the ideal L2 self (e.g., "I can imagine myself as someone who is able to speak Russian") aligned with the immersive aspect of the program and its language pledge to speak only Russian during the summer program (for more information, see Merrill, 2020). Even though the 2021 program was conducted online due to the COVID-19 pandemic and used an updated/modified version of the language pledge (to speak Russian as much as possible), this immersive aspect of the program helped the participants at all initial proficiency levels to increase the clarity of their ideal L2 selves and was particularly beneficial for Curricular Level 3 who had prior L2 Russian learning experience. Moreover, the participants viewed their experiences in the summer program (and before at their prior institutions) positively, as measured by the learning experience part of the L2MSS survey (i.e., the "Cinderella" of the L2MSS framework, see Dörnyei, 2019). Even though learning experience was not always included in L2MSS surveys (e.g., Serafini, 2017), and as a part of a motivation survey/interview needs further refinement, this dimension of L2MSS was beneficial and helped explain the other results, particularly with regards to the ideal L2 self (the learning experience and the ideal L2 self were the strongest components of the participants' L2MSS). The positive view of an immersive Russian learning experience was comparable in certain ways to a positive view of a study abroad experience of L2 Russian

learners (see Zaykovskaya, 2017) or living in other types of "foreign language housing" (see Martinsen et al., 2011).

Another important result with regards to the ideal L2 self was found in the qualitative data, further emphasizing the benefits of using a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018). Students at different stages of their Russian L2 learning had similar visions of their ideal L2 selves: "I do envision myself as someone who will speak Russian in the future" (First-Year Russian or Level 1) and "I definitely vision myself as a person that will speak Russian" (Fourth-Year Russian or Level 4). This finding also illustrated that the students in this summer program were highly motivated L2 learners with clear visions and goals, as well as the financial resources, among others, to support their goals (the summer program is relatively expensive). Considering that this community of students would generally fall under the category of Western, Educated, Industrialized, Rich, and Democratic ("WEIRD," see Henrich et al, 2010) learners who are often referred in SLA literature as "convenient" samples (e.g., Adringa & Godfroid, 2020), this chapter also talks about the importance of collecting data with "non-WEIRD" learners.

In the present study, the scores for the strong ideal L2 self and the weak ought-to L2 self were similar to the ones indicated by Katia, one of the two focal participants in the case study by Pastushenkov and McIntyre (2020). The anti-ought L2 self, however, was prominent for Katia, whereas for the four curricular levels combined this type of self visions was relatively moderate (even though the scores varied from low to high as it was in the analyses at the individual curricular level). Considering that DST scholars have argued that mean scores do now necessarily reflect how the complex dynamic system operate (e.g., Larsen-Freeman, 2014), some of the L2MSS patterns for Katia with regards to the anti-ought-to L2 self (e.g., Thompson, 2017,

2021) were also shown in some of the participants in the present study. It is important to note, however, that the overall low to moderate scores for the ought-to L2 selves and the anti-ought-to L2 selves of the participants (even though the analyses at the individual participant level showed that these scores were relatively high for some of the participants) demonstrated that the students' parents, friends, colleagues, or professors generally supported their decision to study Russian. Despite the trends identified at the four curricular levels combined (further emphasizing the need to look into individual cases, e.g., Papi & Hiver, 2020; Pastushenkov & McIntyre, 2020), some of the participants mentioned that they were told to choose "something more practical like Spanish or French" [Excerpt 4, Participant L, Level 4, Anti-Ought-To L2 Self] or that learning Russian "is a weird thing" [Excerpt 5, Participant D, Level 1, Anti-Ought-To L2 Self] and is "a little bit less important now than it would have been a few years ago" [Excerpt 6, Participant C, Level 2, Anti-Ought-To L2 Self]. Even though these recommendations, on average, did not transform into strong visions of ought-to L2 self or anti-ought-to L2 self, the interview data showed that recommendations not to study Russian existed and that Russian programs may have to address them, further emphasizing the benefits of adding qualitative components to a predominantly quantitative area of ISLA research of IDs (e.g., Serafini, 2017; Thompson, 2017). Adding anti-ought-to L2 self was also important because the field of L2 motivation is dominated by more commonly taught languages (Boo et al., 2015) and this alternative self helps describe the effort or strength against external pressures that L2 learners studying LCTLs such as Russian often experience (see, e.g., Lanvers, 2016; Thompson, 2017, 2021), which was also supported by the qualitative data.

Finally, it is important to note that using two motivational frameworks (L2MSS and General Motivation) helped obtain a more fine-grained view of the dynamics of this ID (see

Serafini, 2017). The General Motivation survey adapted from the works of Gardner and colleagues (e.g., Gardner, 2004) helped assess motivation "right now" at five stages and has been viewed as a more dynamic layer of motivation, whereas Dörnyei's (2009a) L2MSS survey complemented by the anti-ought-to L2 self (Pastushenkov & McIntyre, 2020; Lanvers, 2016; Thompson, 2017, 2021) helped assess how the program changed the participants' visions of their selves and compare how the participants viewed their L2 Russian experiences at their prior institutions and in the summer program (both experiences were generally positive). This dynamic view of L2 motivation as a multi-layered construct aligned with the principles of the complexity/DST framework (e.g., Larsen-Freeman, 2014; Papi & Hiver, 2020). The procedure adopted in this dissertation research helped address the suggestions of Larsen-Freeman (2014) with regards to L2 motivation research guided by DST. In line with Larsen-Freeman (2014), theoretically and methodologically speaking, nothing in the complex system of L2 motivation of L2 Russian learners in the present study was considered fixed (thus the L2MSS data were collected twice, and the General Motivation data were collected five times). These considerations also aligned with Larsen-Freeman's proposition regarding time and space that were at the foremost of the complexity/DST's agenda. However, as indicated by this study's results, the participants' motivation has shown both stable and dynamic patterns. A reasonable criticism of DST in this regard was that it was not sufficient to view IDs as simply 'non-fixed' constructs; it was important to view IDs as 'non-fixed' constructs over time (both longitudinally and at different initial L2 proficiency levels). In other words, this study has shown that it is important to go beyond the idea of nothing is fixed and think about when nothing is fixed. For example, a dynamic pattern for the ideal L2 self was observed at Curricular Level 3 between Phases 1 and 5 (an eight-week difference). The novel behavior of a motivational system emerged through the

interaction of its components (i.e., the consistently high motivation was directly related to the program's effective L2 pedagogy, its language pledge, the immersive aspect, the supportive online instruction due to the COVID-19 pandemic, etc.). The complex system of L2 motivation was comprised of interrelated factors (eight motivational variables and two composite scores were investigated). The system's movement through space and time was non-linear and difficult to predict; it was described in detail in Chapters 6 and 7. Moreover, the system depended on its initial condition or 'the butterfly effect.' An example of this would be the participants' positive learning experience at their prior institutions and the general support from their parents, friends, colleagues, or professors that resulted in a positive and productive learning experience in the summer program. The motivational complex system remained open while interacting with the environment and was feedback- and context-dependent (the program's dedication, support, and pedagogy were crucial). Finally, this study showed that relying only on an average behavior was not (always) informative, as some of the most important patterns discussed earlier in this chapter were found at the individual participant level through the analysis of the quantitative and qualitative data.

Discussion of the dynamics of L2 exposure

In this study, I also investigated the dynamics of the participants' L2 Russian exposure over the course of the eight-week summer program. In this component of the dissertation, I addressed the need to quantify exposure (e.g., Freed et al., 2004; Hanzawa, 2021; Isabelli-García & Lacorte, 2016; Kim et al., 2022; Muñoz, 2008, 2014; Saito, 2015) and track its dynamics at multiple data points (Ranta & Meckelborg, 2013), which is consistent with the principles of the complexity/DST framework (e.g., de Bot et al., 2007; Ellis, 2007; Hiver, 2015; Serafini, 2017). In line with previous research (De Wilde et al. 2022; Leona et al., 2021; Martinsen et al., 2011;

Muñoz, 2014; Saito & Hanzawa, 2018), I also considered the distinction between formal and informal types of exposure. According to the program's description on their official website, each curricular level (1-4) in the eight-week summer program was valued at 9 credit hours and participants met for at least 15 synchronous hours per week. This immersive aspect of the program gave participants approximately 120 hours of formal L2 Russian exposure from the end of June to mid-August in the summer of 2021. As I learned from the dissertation pilot study conducted in the summer of 2019 (Pastushenkov, 2020), students in the summer program were also actively engaged in learning Russian outside their formal classes. Considering that the amount of formal classroom instruction was fixed, I focused on the informal exposure using an L2 exposure log, which has been a commonly used instrument in this type of research (e.g., Martinsen et al., 2011; Pastushenkov, 2020; Ranta & Meckelborg, 2013). The items on the additional L2 exposure log were adapted from the study Ranta and Meckelborg (2013). The log included questions about the students' social interactions, academic work, and recreation. The 13 items were based on the pilot study (Pastushenkov, 2020), my consultations with the program director and faculty, and the analysis of the students' schedule in the summer program. In line with the study by Martinsen et al. (2011), my goal with the additional L2 exposure log was to align the items with the program's co-curricular activities (e.g., informal conversations with peers and tutoring sessions) and popular activities that students were known to be engaged in (e.g., watching movies in Russian, listening to Russian music, and using language learning apps). Remaining consistent with DST principles (e.g., Waninge et al., 2014), it was important to consider the type of the instructed context (Freed et al., 2004; Isabelli-García & Lacorte, 2016; Saito, 2015), as well as the individuality of each learner (Ranta & Meckelborg, 2013). The log was distributed during Phases 2, 3, 4, and 5, corresponding to Weeks 2, 4, 6, and 8 in the summer program. Even though it would have been beneficial to distribute the log more frequently, for some of the participants it was difficult to find time to participate in research and I needed to find a balance between the number of phases and the students' availability.

The preliminary analyses of the quantitative data from the additional L2 Russian exposure log showed that one of the items (Item 5 - Homework in the Russian program) was both quantitatively and qualitatively different from the other 12 items. As for the quantitative differences, homework during the Russian program was the most time-consuming activity compared to the other categories. On average, the participants reported that they spent 14 hours on homework from Monday through Friday during Weeks 2 and 4 and 11 hours during Weeks 6 and 8. In terms of the qualitative differences, homework was also directly linked to the participants' formal exposure. Upon my consultation with the program director and faculty, I decided to exclude "Item 5 - Homework in the Russian program" from further analyses and focused on the remaining 12 activities. However, it is still important to note that in addition to 120 hours of formal instruction, the participants spent at least 120 hours on average on homework in the summer program, which gave them approximately 240 hours of more formal types of exposure during the summer program.

Similarly to the motivation data analysis discussed earlier in this chapter, I explored the students' additional L2 exposure at the four levels combined, separately at the curricular levels, and also individually at the participant level. Analyzing data at the individual participant level also aligned with the principles of the complexity/DST framework (e.g., Papi & Hiver, 2020; Larsen-Freeman, 2014) and helped identify some of the trends in the data, as well as one of the study's limitations (some of the students may have overestimated the hours they spent on learning Russian informally). The general pattern of stability of motivation at the four curricular

levels combined was also observed in the quantitative data from the additional L2 exposure log. Overall, the participants were consistently and actively engaged in learning Russian outside their formal classes in the summer program. Even though the quantitative data from the log indicated that the participants' additional L2 Russian exposure in mean total hours/week slightly decreased from Phase 1 (M = 16.95, SD = 9.90, 95% CI [13.38, 20.52]) to Phase 5 (M = 14.28, SD = 9.48, 95% CI [10.18, 18.38]), the 95% CI for the four phases overlapped, suggesting that there were no statistically significant changes in the amount of additional L2 exposure that the students at the four curricular levels combined were engaged in. With some minor variation, the same pattern of stability was observed at the individual curricular levels and at the participant level. It is important to note, however, that the students' total number of hours that they spent learning Russian outside their formal classes was higher on average at Curricular Levels 2 and 3 when compared to Level 1. Despite these nuances, the students at all levels actively supported their immersive experience by using social media/email; talking to their instructors, tutors, and peers in Russian outside their formal classes; using language learning apps; using Russian while playing sports or video games; reading Russian literature, news, and blogs; watching TV/movies/streaming services in Russian; listening to music in Russian; using Russian while participating in thematic clubs; and attending Russian guest lectures. These findings regarding the students' active engagement in learning a L2 informally were consistent with previous research conducted in foreign language housing (see Martinsen et al., 2011), a context that is similar to the Russian summer immersion program. Overall, I estimated that the students spent at least 120 hours, on average, learning Russian outside their formal classes. Combined with the formal exposure from classes and homework (240 hours), the students spent approximately 360 hours learning Russian during the summer program. This high level of both formal and informal

exposure transformed into considerable learning gains that were discussed earlier in this chapter. These results were comparable to one year of university instruction (see, e.g., Merrill et al., 2021; Rifkin, 2005). It is important to note that not only the quantity of L2 exposure may have played a role, but also the quality of L2 exposure, as Muñoz pointed out that "intensive exposure seems to be more effective than long periods of drip-feed instruction" (p. 16). Overall, remaining consistent with previous propositions (see, e.g., Muñoz, 2014; Ranta & Meckelborg, 2013; Saito & Hanzawa, 2018), this study showed that high levels of high-quality L2 exposure resulted in considerable learning gains.

The analysis of the quantitative data revealed that the two most popular activities that the students were engaged in were: listening to music in Russian and watching TV/movies/streaming services in Russian (on average, 3 hours in Week 2, 4, and 6 in the two categories). During Week 2, watching TV/movies/streaming services in Russian was the most popular activity (also, 3 hours on average), while the second most popular activity was conversations with peers in Russian (2 hours on average). During Weeks 2, 4, and 6, listening to music and watching TV/movies/streaming services were followed by conversations with peers in Russian and conversations with instructors in Russian. On average, the participants spent 2 hours on these conversational activities during Week 2, 4, 6, and 8. During Week 8, language learning apps were the fourth most popular activity (for more information about app use by L2 Russian learners, see Pastushenkov, 2020). Overall, the analysis of the quantitative data from the additional L2 exposure log showed that the participants had well-established learning routines and spent approximately the same amounts of time for each of the learning activities outside their formal classes over the course of the summer program. It is important to note, however, that the students listened to music and participated in thematic clubs less by the end of the program,

which may have been indicative of the fatigue that they experienced after a summer of intensive L2 Russian learning. The students' participation in tutoring sessions also fluctuated, which may have to do with the fact that the Russian instructors asked their students to attend tutoring more by the end of the program.

The general trends identified in the quantitative data were supported by the interview data, which was not included in some of the previous studies of L2 exposure (e.g., Muñoz, 2014; Pastushenkov, 2020; Ranta & Meckelborg, 2013). From the interviews, I learned that some of the participants were engaged in learning Russian outside the summer program because they tried to compensate for the lack of face-to-face immersion due to the COVID-19 pandemic: "I just tried to immerse myself as much as I could because I wasn't there in-person on campus" [Excerpt 19, Participant V, Level 1, Additional L2 Exposure/Learning Experience]. The qualitative data helped figure out some details about how the participants learned Russian outside their formal classes. For example, one of the participants said: "I watched a lot of Russian YouTube videos just to hear other people speak it and ... there's this channel called "Easy Russian" and they just interview people on the streets, and I watched that a lot" [Excerpt 20, Participant A, Level 1, L2 Additional L2 Exposure]. Overall, using a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018), helped obtain a more fine-grained understanding of the participants' engagement in L2 exposure and its dynamics during the summer program.

Discussion of the role of IDs in L2 Russian learning

In this study, I also investigated how L2 Russian students' WM capacity, motivation, and the amount of L2 exposure affected their learning gains in lexicogrammar, speaking, and writing. The learning gains were calculated by subtracting the students' placement scores from their exit scores on the four tests (lexicogrammar FIB quiz, lexicogrammar MC quiz, ACTFL OPIcs, and ACTFL WPTs). The inclusion of these conceptually different IDs (cognitive, affective, and exposure-related factors) and different language skills, as well as adding a longitudinal perspective and multiple phases, was guided by the principles of the complexity/DST framework (e.g., de Bot et al., 2007; Ellis, 2007; Larsen-Freeman, 2014; Serafini, 2017; Waninge et al., 2014) and multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013). In line with Serafini (2017) and other SLA researchers (e.g., Dörnyei, 2010a; Dörnyei & Ryan, 2015; Larsen-Freeman, 2014; Segalowitz & Trofimovich, 2012; Toth & Davin, 2016), this study supported the need to re-evaluate problematic dichotomies such as the cognitive/affective dichotomy.

The quantitative results showed multiple instances when a higher WM capacity played a facilitative role in the L2 Russian learning. The learning gains in L2 Russian writing correlated with the scores on both the OSpan (r = .94, p < .001, 95% CI [.84, .99]) and the BDS tasks (r = .62, p = .06, 95% CI [.38, .92]) at Curricular Level 3. Another strong correlation was also found at Level 3 between the writing gains and the scores on the BDS task (r = .62, p = .06, 95% CI [.38, .92]). The 95% CI for the effect sizes supported the facilitative role of WM in the L2 writing gains at Curricular Level 3. The composite score of WM was a statistically significant predictor of the FIB exit scores and the MC exit scores. The regression and correlation analyses also showed some evidence that WM played a stronger role at Levels 1 and 3. Sagarra (2017) pointed out that more robust WM effects have been found in beginner learners (e.g., Linck & Weiss, 2011, 2015; Serafini & Sanz, 2016) than in more advanced students (e.g., Frost et al., 2013; Grey et al., 2015), whereas this study showed that stronger WM effects were found at both the beginner and more advanced levels, but not at the intermediate level. It is important to note

that another research study conducted in the same Russian immersion program (Kisselev et al., forthcoming) showed that the complexity of lexicogrammar was a strong predictor of the students' L2 writing scores. In line with these findings, a higher WM capacity was associated with gains on lexicogrammar-centered tests (FIB, MC, and WPT), but did not explain the discrepancies in the gains on the ACTFL OPIcs. The results of this study are consistent with the meta-analysis by Linck et al. (2014) focusing on studies in instructed settings that showed that a higher WM capacity positively affected proficiency outcomes (an estimated population effect size (ρ) = .255). It is also important to note, however, that the sample size in this study was low and thus these results need to be interpreted with caution.

The interviewees also acknowledged that memory was a variable that they considered throughout their L2 Russian studies in the summer program and that many of them tried to develop strategies or routines that would help them memorize Russian vocabulary and grammar more efficiently. Even though these strategies may be related to both WM/short-term memory and long-term memory abilities, the role of WM in L2 learning of vocabulary and grammar, particularly the fact that WM capacity is limited, has also been discussed in previous literature about memorization strategies (e.g., multimedia-assisted self-learning materials and digital flashcards, Zhu et al., 2012; Li & Tong, 2019; color coding, Ibarra Santacruz, & Martínez Ortega, 2018; mnemonic strategies, Di Santo et al., 2020). One of the participants said that "having online material was a little harder for me... I found that when I was rewriting my exercises, I would remember the vocabulary and the grammar rules much better" [Excerpt 28, Participant R, Level 2, WM/Additional L2 Exposure]. Another student mentioned that it was easier to memorize new Russian words and grammar rules before going to sleep. The participants also talked about the importance of flashcards and other notes: "I found that the most

successful thing for me to do is a lot of note cards, and I have to repeat them... I hear them and then I repeat them sometimes... in the mornings when I go to the gym, the first 20 minutes, I review my notes" [Excerpt 29, Participant S, Level 3, WM/Additional L2 Exposure].

Overall, the relationships between motivation and learning gains were less straightforward than the positive correlations between WM capacity, additional L2 exposure, and the learning gains. A potential reason for that was that the participants were generally highly motivated L2 Russian learners when they joined the program. The analyses of the quantitative data also showed that the participants' vision of the ideal L2 self has changed significantly at Curricular Level 3. The analyses at the individual participant level also showed that there were instances that the participants' motivation assessed by the L2MSS and General Motivation surveys decreased (even though on average, the motivation scores went up). In the bivariate correlation analyses, I identified 10 large and statistically significant effects (positive and negative). Most of these effects were not considered representative of the population due to the broad confidence intervals. Strong negative correlations in Phase 1 were found between ought-to L2 self and the gains on the lexicogrammar FIB quiz (r = -.75, p < .001, 95% CI [-.93, -.31]) and the writing test (r = -.69, p = .03, 95% CI [-.92, -.10]) at Level 3. In other words, strong external pressures (from peers, teachers, parents, etc.) encapsulated by the ought-to L2 self (Dörnyei, 2010; Thompson, 2017) were not associated with higher learning gains. Previous research with L2 Russian learners also showed that a stronger vision of the ought-to L2 self, which is associated with extrinsic motivation (Lai, 2013), did not necessarily transform into considerable gains and long-term commitment to learning Russian (Pastushenkov & McIntyre, 2020). Ushioda and Dörnyei (2012) pointed out, however, that if L2 proficiency is a part of an individual's vision of ideal L2 self or ought-to L2 self, this individual will be motivated to

acquire the L2. In this dissertation, a stronger ought-to L2 self was not associated with higher gains, except for the strong positive correlation between ought-to L2 self and the gains on the OPIc was found in Phase 5 at Curricular Level 1 (r = .84, p = .04, 95% CI [.09, .98]). It this case, the 95% CI were broad, and this correlation was not considered representative of the population.

A rather surprising result was the negative correlation in Phase 1 at Curricular Level 1 between General Motivation and the gains on the speaking OPIcs (r = -.72, p = .01, 95% CI [-.92, -.24). From a theoretical standpoint, it is difficult to explain this result because motivation has been generally viewed as vehicle for L2 learning (e.g., Sato & Loewen, 2019a). A potential explanation is that at early stages of L2 development, which was the case for Curricular Level 1, the relationships between motivation and learning gains were more difficult to predict, which DST proponents would refer to as erratic behaviors as opposed to more stable behaviors referred to as attractor states (e.g., Hiver, 2015; Larsen-Freeman, 2012; Papi & Hiver, 2020; Serafini, 2017). At Curricular Level 2 (or at later stages of L2 development), General Motivation was positively correlated with the gains on the speaking OPIcs in Phase 2 (r = .66, p = .01, 95% CI [.17, .89]) and Phase 3 (r = .72, p = .03, 95% CI [.11, .94]). However, erratic patterns were also observed at Curricular Level 3: ideal L2 self was negatively correlated with the OPIc gains (r = -.95, p <.001, 95% CI [-.99, -.69]); ought-to L2 self was negatively correlated with the gains on the FIB quiz (r = -.91, p = .01, 95% CI [-.99, -.48]). Interestingly, General Motivation in Phase 5 (at the very end of the summer program) was a marginally significant predictor of the MC exit scores. Due to the low sample size (particularly in Phase 5), the results of the regression analyses need to be interpreted with additional caution. Overall, more research in the area of L2 motivation in instructed L2 learning of Russian is needed, including with students outside the

unique context that is the Russian summer immersion program (i.e., a "WEIRD" context, see, e.g., Adringa & Godfroid, 2020; Henrich et al, 2010).

The analysis of the bi-weekly log data showed several instances when learning Russian outside formal classes played a facilitative role in the learning gains at different curricular levels, particularly later in the program in Phases 4 and 5. For example, the additional L2 exposure in Phase 4 correlated with the students' learning gains on the lexicogrammar FIB quiz at Level 3 (r = .63, p = .13, 95% CI [.07, .96]). However, none of these effects are considered representative of the population due to the broad confidence intervals. This finding was also supported in the multiple linear regression analyses for Phases 2-5. Additional L2 exposure was not a statistically significant predictor of the exit scores and did not add value to the regression models. This finding, however, does not undermine the facilitative role of exposure in instructed L2 learning (see, e.g., Freed et al., 2004; Isabelli-García & Lacorte, 2016; Kim et al., 2022; Muñoz, 2014; Saito, 2015). A potential explanation for the fact that additional L2 exposure was not predictive of learning gains was that the students, as it was pointed out earlier in this chapter, were all actively and consistently engaged in learning Russian outside their formal classes. Methodologically speaking, it would have been beneficial to have both? students who were engaged in additional L2 exposure and students who were not engaged in this type of learning. However, this is not how instructed L2 learning works in ecologically valid settings. Another possible explanation would be the tests used by the program. For example, additional L2 exposure may have helped the participants learn more about Russian culture, the knowledge of which was not tested in the placement and exit tests. Overall, the interviewees from all curricular levels tended to find their learning activities outside the formal classes beneficial for their

studies. In line with previous research with L2 Russian learners (Zaykovskaya et al., 2017), I followed the principle of trusting the interviewees.

Theoretical implications: DST and future ISLA/ID research

The study's entirety, including its design and analyses, was guided by DST (e.g., de Bot et al., 2007; Ellis, 2007; Hiver, 2015; Larsen-Freeman, 1997, 2012, 2014; Serafini, 2017; Papi & Hiver, 2020; van Geert, 1994; 1995; Waninge et al., 2014) and multidimensional perspectives on IDs in instructed L2 learning (e.g., Sun & Zhang, 2020; Wang & Eccles, 2013). This theoretical orientation was reflected in the study's focus on the three conceptually different ID categories (WM, motivation, and L2 exposure) and different language skills (lexicogrammar, speaking, and writing); its longitudinal design with multiple phases and different initial proficiency levels; the way the context and individual participants were considered in the development of the study's instruments (particularly, the two motivation surveys and the additional L2 exposure log); as well as in the analyses that were conducted for the all levels combined, separately at different curricular levels, and individually at the participant level (with the qualitative data from the interviews being a particularly helpful addition). This study was theoretically and methodologically motivated by the study by Serafini (2017), who referred to other proponents of DST working in the areas of SLA and ISLA and argued that a dynamic systems approach was "the logical next step of conceptualizing IDs" (Dörnyei, 2010a, p. 260) even though "it is humanly impossible to study everything at once" (Larsen-Freeman, 2012, p. 208).

Despite its untapped potential (Hiver & Al-Hoorie, 2016), empirical ISLA studies guided by DST, including this dissertation research, can be challenging to prepare and conduct. In general, this type of research would benefit from larger samples, which is difficult to implement in a study of LCTLs such as Russian, as well as adding more phases to the study's design, which

can also be challenging because students do not always have time (or do not want) to participate in research. Even though this study is considered longitudinal, conducting an empirical ISLA study over the course of several years would be beneficial for the field. Alternatively, L2 researchers can adopt retrospective qualitative narratives, even though this methodology also has its limitations (see, e.g., Papi & Hiver, 2020; Pastushenkov & McIntyre, 2020). Moreover, empirical ISLA studies of IDs guided by the complexity/DST framework should consider using more advanced statistical analyses such as regression models with interaction terms (see Figures 33 and 34) and structural equation modelling (see Figure 35), which would also require larger samples (see, e.g., DeKeyser, 2012; Hayes, 2017; Winke 2013).

Statistically speaking, an exploration of compensatory relationships between IDs (one of the future directions of this project) will require regression-based mediation and moderation analysis (for review, see Hayes, 2017). An example of a mediation model is shown in Figure 33 and an example of a moderation model is shown in Figure 34. An "intervening variable," a mediator, is conceptualized as a mechanism through which one variable influences another variable, while a moderator, also known as interaction, affects a relationship between two variables in terms of its size or sign (Hayes, 2017). Hayes uses previous research on how media affects one's perceptions of one's own physical shape to illustrate the differences between the two concepts. An example of mediation would be media causing dissatisfaction about one's physical shape for people who have a weaker internalized vision of an ideal body shape; and an example of moderation would media intensifying the effect of dissatisfaction about one's physical shape for people who already have an internalized vision of an ideal body shape. DeKeyser (2012) pointed out that "interactions between individual variables and external, educational or contextual variables allow for more fine-tuned (and hence more generalizable)

predictions that help with adaptation of teaching methodologies to students or matching students with treatments" (p. 190).



Figure 33. Simple mediation model with one mediator (Hayes, 2017).



Figure 34. Simple moderation model with one moderator (Hayes, 2017).



Note. All values adjacent to arrows are standardized estimates. Single-headed arrows are direct effects. Double-headed arrows are correlations. e = error variance.
*Standardized coefficient p < .05.
*Standardized coefficient p < .01.

Figure 35. Structural Equation Model with Parameter Estimates (Winke, 2013).

The materials, procedure, and analyses used in this study followed previous ID research guided by DST (e.g., Serafini, 2017), as well as quantitative (e.g., Kisselev et al., forthcoming), qualitative (e.g., Zaykovskaya, et al., 2017), and mixed-methods (e.g., Pastushenkov & McIntyre, 2020) studies with L2 Russian learners. In this empirical ISLA study guided by the complexity/DST framework, it was important to consider different perspectives and research methodologies; reconsider problematic dichotomies (e.g., the cognitive/affective dichotomy, see, e.g., Dörnyei, 2010; Serafini, 2017); as well as use different ID frameworks such as Gardner and colleagues' AMTB (Gardner et al., 2004; Gardner & Lambert, 1972; Masgoret & Gardner, 2003) and Dörnyei's (2009) L2MSS complemented by the newly emerged anti-ought-to L2 self (Lanvers, 2016; Pastushenkov & McIntyre, 2020; Thompson, 2017, 2021). Even though it is impossible to consider everything at once (Larsen-Freeman, 2012), looking for new ID links such as the cognitive/affective/exposure link was also beneficial, considering that these IDs have

been previously investigated together (e.g., WM and motivation see Serafini, 2017; Winke, 2013; and WM and L2 exposure, see Denhovska et al., 2016; Indrarathne, & Kormos, 2018). As various SLA scholars have pointed out (e.g., de Bot et al., 2007; Dörnyei, 2009a; Sparks et al., 2012; Pawlak, 2012; Sun, 2019), IDs are interrelated and they engage in complex relationships that fluctuate over time, which is one of the key principles of the complexity/DST framework (de Bot et al., 2007; Ellis, 2007; Larsen-Freeman, 2014).

In preparation of an empirical ISLA study of IDs guided by DST, which was also a timeconsuming process as shown in this dissertation research (for more detail, see Chapters 3 and 4), as Serafini (2017) pointed out, L2 researchers may consider Hiver and Al-Hoorie's (2016) "dynamic ensemble" (i.e., practical suggestions for L2 research guided by this framework). Hiver and Al-Hoorie suggested that future studies will need to consider operational factors (e.g., "What is the complex system under investigation?"), the role of context (e.g., "What are the contextual factors that are part of the environmental frame of reference for the system, its dynamic actions, and its patterned outcomes?"), as well as consider both the macro-system level (e.g., "What general principles of change exist for this system?") and its micro-structure (e.g., "What are the parts that make up the system under investigation?"). Despite its limitations, particularly its limited sample size and the use of less sophisticated statistical methods, this study has taken these factors into consideration, which helped shed some light on the role of IDs and their dynamics in instructed L2 learning of Russian, an area that remains understudied within the field of ISLA (Kisselev et al., forthcoming; Merrill et al., 2021).

Mixed-methods research and LCTLs

Despite the clear need to investigate the SLA and ISLA of LCTLs and critical languages (e.g., Kim, 2017; Pastushenkov & McIntyre, 2020; Thompson 2017), Russian and many other

languages remain underrepresented in these domains. This project has the potential to spark interest in this area and outline some future directions for ISLA research with LCTLs. In addition to using DST as a theoretical framework, one of such future directions is the use of the mixedmethods approach, including a mixed-methods sequential explanatory design (see, e.g., Ivankova et al., 2006; Taguchi, 2018). Even though this methodology can be challenging to prepare and implement (for more detail, see Chapters 3, 4, and 5), this approach helped better understand some of the findings from the quantitative ID data. For example, the interviews with Participants M and V were consistent with the quantitative data from the two motivation surveys (L2MSS and General Motivation). Participant M stated that their "motivation would turn from simply maintaining a working proficiency to now I want to like see the world with the Russian eyes" [Excerpt 1, Participant M, Level 4, General Motivation/Ideal L2 Self], which is also consistent with the significant changes in the vision of the ideal L2 self at more advanced proficiency levels. Participant M mentioned that their "motivation to get back into Russian was really high in the beginning and then throughout the course it got even higher because of how challenging the language is" [Excerpt 2, Participant V, Level 1, General Motivation], remaining consistent with the general trends identified in the quantitative data.

This mixed-methods sequential explanatory design was also beneficial because of the low sample size in this dissertation research, which is a common issue for ISLA studies of LCTLs (Thompson 2017), as well as the field of ISLA in general (Loewen & Hui, 2021). Considering that the results of statistical analyses with low sample sizes need to be interpreted with additional caution, adding a qualitative component helped support the findings and learn new information about the participants. For example, I learned about some of the participants' memorization strategies (e.g., using flashcards, re-writing notes, using a mix of Duolingo and Russian movies,

reviewing new words while attending a gym, learning new Russian words before going to sleep, etc.), as well as about some of the resources that the participants used outside their formal classes (e.g., the YouTube channel called "Easy Russian" or the podcast called "Russian Progress" with Artem Nazarov).

In some cases, the interviewees provided information that was not observed in the quantitative data, even though these insights were important from administrative and pedagogical standpoints. For example, the scores for the participants' ought-to L2 self and anti-ought-to L2 self were generally low or moderate. However, I learned from the interviews that they were cases when parents, friends, or teachers discouraged the students from learning Russian and recommended "something more practical like Spanish or French" [Excerpt 4, Participant L, Level 4, Anti-Ought-To L2 Self] or even said that learning Russian "is a weird thing" [Excerpt 5, Participant D, Level 1, Anti-Ought-To L2 Self]. These insights provided by the qualitative data have direct implications for language programs who may need to explain the benefits of learning Russian and share successful and unsuccessful stories of their graduates, which was also pointed out in the mixed-methods case study by Pastushenkov and McIntyre (2020).

As opposed to the generally quantitative domain of cognitive ISLA research, studies of L2 motivation have been quantitative (e.g., Serafini, 2017; Thompson, 2017), qualitative (e.g., Papi & Hiver, 2020; Thompson & Vásquez, 2015; Sun & Zhang, 2020) and mixed-methods (e.g., Fathi et al., 2019; Pastushenkov & McIntyre, 2020). The mixed-methods approach has been considered a promising methodology in ID research, even though Boo et al. (2015), when speaking of L2 motivation research, pointed out that "most of the current practice seems to involve rather superficial mixing of relatively independent qualitative and quantitative

components within a study" (p. 156). Therefore, more robust mixed-methods studies of IDs are needed in the field of ISLA, including in the domain of ISLA studies with LCTLs.

Ecological validity and ISLA research with LCTLs

Considering that previous SLA and ISLA research was dominated by studies of more commonly taught languages such as English (e.g., Foster et al., 2014; Leona et al., 2021; Papi & Hiver, 2020) and Spanish (e.g., Grey et al., 2015; Sagarra, 2017; Serafini, 2017), I also strove to promote the need for more ecologically valid research with LCLTs such as Russian. Conducting this research at the eight-week Russian summer immersion program was a valuable and rare opportunity that also extended previous laboratory-based studies in the area of Russian ISLA to an authentic setting (e.g., Denhovska, et al. 2016). Ecologically valid research is beneficial from ISLA perspectives because it can help promote the research-pedagogy dialog (see, e.g., Sato et al., 2021). This dissertation research, for example, illustrated that L2 learners' IDs not only helped explain some of the discrepancies in their learning gains (specifically on the lexicogrammar-centered FIB quiz, MC quiz, and WPT), but also that the analysis of the dynamics of motivation can serve as a proxy of the effectiveness of the program's pedagogy, which will be discussed later in this chapter. Even though conducting research in ecologically valid settings with natural languages and real L2 learners has its challenges, as shown in this dissertation research (for more detail, see Chapter 4), as Godfroid (2016) pointed out, "... learning does not happen in a vacuum" (p. 207).

Ecologically valid research has the potential to address some of the criticisms of the field of SLA with regards to its relevance and applicability to L2 pedagogy (see Marsden & Kasprowicz, 2017; Sato & Loewen, 2019b; Spada, 2015). In line with previous propositions (e.g., Sagarra, 2017; Serafini, 2017), I considered various factors in this dissertation research,

including the L2 learners' backgrounds and their initial L2 proficiencies. To the best of my knowledge, this is one of the few studies in the domain of L2 Russian SLA/ISLA that considered L2 proficiency as an important variable (see, e.g., Cho & Slabakova, 2014). Despite its challenges, conducting this study at a major Russian immersion program has the potential to spark more interest in the understudied domain of ecologically valid ISLA research with LCTLs such as Russian (see, e.g., Kisselev et al., forthcoming).

Collecting data online during the COVID-19 pandemic

Due to the COVID-19 pandemic, the summer program, as well as this dissertation research, were conducted entirely online. In this study, I used Qualtrics for the L2MSS and General Motivation surveys and the additional L2 exposure log; Gorilla Experiment Builder for the WM tasks; and Zoom for the interviews. Collecting data online had its benefits and disadvantages. I generally viewed Qualtrics, Gorilla, and Zoom as convenient tools for data collection, even though collecting data online as opposed to working with L2 learners in a language classroom or a research laboratory gave me less control over this process. Due to this issue, I implemented several steps of data cleaning described in Chapter 5. The raw data exported from Qualtrics and Gorilla also included the time that the participants spent on the surveys and tasks, as well as showed their completion rate, which helped clean the data. In terms of the data cleaning and data processing in general, I tried to follow recommendations outlined in previous literature (e.g., Kisselev et al., forthcoming; Lodder, 2013; Sainani, 2015; Serafini, 2017; Serafini & Sanz, 2016; Zaykovskaya et al., 2017), including the studies that were also conducted online using online tools like Qualtrics (e.g., Sato et al., 2021).

It is also important to note that I was originally supposed to collect data for this dissertation in the summer of 2020, after the pilot study that I conducted in the summer of 2019

(Pastushenkov, 2020). Due to the COVID-19 pandemic, the program had to switch to an emergency remote model of instruction. The number of students in the online program of 2020 was also lower than in previous years (83 students total). Considering these circumstances, I decided to postpone my data collection to the summer of 2021, which also gave me more time to modify the materials and procedure (for more detail about the modifications, see Chapter 4). The summer program of 2021 was also conducted entirely online; however, the instructional model was not considered emergency remote, and the number of students enrolled in the program was higher (even though the program only included four curricular levels as opposed to seven levels that were included in previous years before the pandemic). Moreover, the enrollment went to 111 students at the four curricular levels in the summer of 2021 compared to 83 students in the summer of 2020.

Despite the effort that the program director and faculty put into its online pedagogy, which was also supported the consistently high motivational patterns of the participants in this dissertation research, some of the students mentioned that having the program conducted entirely online had its limitations. For example, one of the students mentioned that having the program "completely online was definitely a difference ... when I studied Spanish in middle and high school that was all in person, so there were more opportunities to actually practice speaking the language with your peers and to get that aspect of learning the language" [Excerpt 7, Participant V, Level 1, Learning Experience/Additional L2 Exposure]. Even though the students' gains over the course of the summer program were considerable, research in this summer program has shown that lexicogrammar gains in the face-to-to-face program tended to be higher than in the online program, especially at the intermediate level, partially because the face-to-face program provided more opportunities for interaction outside formal classes (see Merrill et al., 2021). In

light of these observations, a possible future direction would be to replicate this dissertation research in a face-to-face Russian summer immersion program.

Practical implications

Pedagogical implications

Despite the fact the program was intensive and conducted entirely online due to the COVID-19 pandemic, the students' motivation was consistently high (or even increased) at the four levels combined, suggesting that the program and the instructors were able to create enriching activities for its students and helped them maintain their motivation over the course of the program. Considering that even the most motivated L2 Russian learners can lose their motivation because of the instruction that they receive (see Pastushenkov & McIntyre), it is important to discuss the philosophy and pedagogy of this summer immersion program. Despite the challenges of online teaching during the pandemic, the program strove to maintain its teaching principles, including the focus on communicative activities and its language pledge to speak only Russian, which was positively viewed by the participants as shown in the interview data. Several interviewees mentioned that this practice was new to them because they were more used to doing Russian grammar drills at their home institutions, which was confirmed by the program director and faculty who have extensive teaching and collaborative experience in the United States. The summer program also focused on work in small groups, while also giving students individualized attention and opportunities for conversations. As I learned from the interviews, such conversations gave students opportunities to practice new vocabulary and grammar, which they also viewed positively. To help give students individualized attention, the program hired a group of tutors, which is not a common practice for most traditional Russian language programs in the United States. The tutors (who were experienced L2 instructors) helped
students outside of their formal classes. This is why this type of exposure was included in the additional L2 exposure log. Even though tutoring was not the most popular activity among the students (for more detail, see the first part of this chapter or Chapter 6), they were consistently engaged in this type of additional L2 exposure over the course of their studies in the summer program. The tutors helped learners even out their language skills in speaking, listening, reading, and writing, as well as fill in the gaps in the knowledge of Russian if necessary, or practice their phonetic skills.

It is important to note that the program was intensive. The curriculum in the online program of 2021 covered approximately one academic year of Russian, which was also reflected in the students' considerable learning gains. In addition to the language pledge and focus on communicative activities, the program put an emphasis on co-curricular and cultural activities. Outside of their formal classes, the students had different types of Russian co-curricular and cultural practice organized by the program: office hours and tutoring; cultural events that included guest lectures, journalism workshops, and thematic clubs on different cultural topics such as Russian cartoons and internet culture; and informal socialization opportunities using Zoom and Discord. Overall, this study showed that the program's pedagogy was effective and resulted in considerable gains, which was particularly difficult to implement during the COVID-19 pandemic. The analysis of the dynamics of motivation served as a proxy of the effectiveness of the program's pedagogy. The motivation surveys (particularly the General Motivation survey) used in this dissertation research can be implemented in other language programs. These instruments were considered reliable and valid L2 motivation instruments (for more detail, see Chapter 5). The General Motivation survey included only five questions and can be complemented by an open-ended question that can ask students to describe the reasons for their

responses (similarly to the interview data from this dissertation research). Instructors of L2 Russian and other languages can distribute this survey at different points through the semester, similarly to this study, which is also in line with the learner-centered approach, a key principle of the summer program's teaching philosophy.

In addition to motivation and L2 exposure, this study investigated the role of WM capacity in instructed L2 learning of Russian. This holistic, non-modular perspective on IDs can also help inform L2 pedagogy as "much about effective pedagogy depends on striking an optimal, Goldilocks-like balance between instructional goals, means, and support" (Toth & Davin, 2016, p. 152). The analyses of the quantitative data revealed several instances when WM played a facilitative role in L2 development, particularly with regards to the lexicogrammar centered FIB quiz, MC quiz, and ACTFL WPT. Considering that it is important to "keep in mind that different [pedagogical] approaches and methods favor learners with different cognitive strengths" (p. 411), Li (2017) provided different tips for teachers in light of previous ISLA research, including the importance to adapt certain aspects of instruction to help L2 learners of various cognitive profiles; use a combination of inductive and explicit tasks (with more explicit focus toward the end); provide salient feedback and consider adding pre- and post-tasks than can help ease cognitive burden; give L2 learners opportunities for planning before and within task performance; and provide support for cognitively demanding tasks (e.g., in a form of scaffolding and linguistic support). The interviewees also acknowledged that memory was a variable that many of them considered during their Russian studies and tried to develop their own memorization strategies, which were also discussed earlier in this chapter. Overall, as shown in this dissertation study, learner-centered ID research can help inform learner-centered pedagogy.

Implications for the development of background questionnaires and placement testing

L2 researchers have previously emphasized that conceptually different categories of IDs can be interconnected (e.g., Dörnyei, 2010; Dörnyei & Ryan, 2015; Larsen-Freeman, 2014; Segalowitz & Trofimovich, 2012), which was realized in empirical ISLA studies at the intersection of cognitive and affective IDs (e.g., Serafini, 2017; Winke, 2013), as well as in research focusing on IDs in cognitive abilities and L2 exposure (e.g., Denhovska et al., 2016, Pattemore & Muñoz, 2020). As Dörnyei (2010) suggested, "the separation of cognition and motivation has been increasingly seen as an outdated and inaccurate conceptualization" (p. 259). Based on the results of this study, emphasizing the need to consider learners' IDs as important elements in the complex system of L2 development, ID research in the domain of ISLA, including this dissertation study, also has important implications for the development of future background questionnaires and placement testing.

Various researchers emphasized the need to consider the role of context (e.g., Dörnyei, 2009a; Faretta-Stutenberg & Morgan-Short, 2018; Larsen-Freeman, 2014; Papi & Hiver, 2020; Pattemore & Muñoz, 2020) and L2 proficiency (e.g., Pastushenkov, 2020; Sagarra, 2017; Serafini, 2017) in ID research, which also aligns with the principles of the complexity/DST framework (e.g., Waninge et al., 2014). For example, Denhovska and colleagues (e.g., Denhovska et al., 2016), in their laboratory-based studies of instructed L2 Russian learning, excluded participants who had prior knowledge of Slavic languages because the researchers wanted to encapsulate learning at the very beginning stages of L2 development. Considering that there is strong empirical evidence from previous ISLA research on the facilitative role of WM, motivation, and L2 exposure in L2 development, future studies in this domain may consider

adding cognitive, affective, and exposure-related IDs in their background questionnaires. For example, Sagarra (2019) pointed out that future ISLA studies would benefit from adding WM in their design. In the era of online data collection tools such as Qualtrics and Gorilla Experiment builder used in this dissertation research, these types of background questionnaires are much easier to distribute.

As I learned from my frequent consultations with the program director and faculty, L2 learners' IDs, particularly their motivation, can also be assessed in future placement testing used in the summer program. In line with previous research (e.g., Pastushenkov & McIntyre date?), this study once again showed that motivation to study Russian may vary. The case of Participant D was very interesting in this regard. This student considered doing a Ph.D. in Soviet history. The participant's advisor suggested that learning Russian would assist in achieving this goal. During the interview, the participant mentioned that "I don't have a lot of intrinsic motivation to learn Russian" [Excerpt 16, Participant D, Level 1, Ought-To L2 Self] and added that "I don't enjoy learning languages as something just to do for fun... I find it really difficult and headache inducing" [Excerpt 17, Participant D, Level 1, General Motivation/Learning Experience]. The motivation of Participant D to learn Russian was conceptually different compared to that of the other interviewees (entirely extrinsic motivation vs a combination of intrinsic and extrinsic factors). Based on these results, future language programs may benefit from asking their students about their motivation to study Russian or other languages during placement testing. For example, students who want to do a Ph.D. in Soviet history may be placed in a separate track for historians, rather than in levels that are based only on the students' L2 proficiency.

Limitations and future directions

The participants in this dissertation research were 52 students recruited from an eightweek Russian summer immersion program from four curricular levels. Despite the fact that 52 out of 111 students initially participated in the study, the number of students who participated in later phases was lower, especially at Curricular Level 4, which was excluded from some parts of the analyses. The issue of participant attrition was evident. Having a limited sample size has been a common issue in ISLA studies of LCTLs (Thompson 2017), including Russian (e.g., Pastushenkov, 2020), as well as in the field of ISLA in general (Loewen & Hui, 2021). As Loewen and Hui pointed out, "researchers inherit the issue of small samples from small L2 classes" (p. 191) and a possible solution to this problem is replication research. In the case of the summer program, some students were too busy to participate in research because the program was so intensive. As a future direction for this project, I consider returning to the face-to-face Russian summer immersion program with a replication study that would also explore the program's pedagogy in more detail (e.g., in the form of interviews with the instructors and classroom observations). It is important to note, however, that conducting a replication study may not address the sample size issue. Therefore, a mixed-methods design, described earlier in this chapter, would be an important future direction for ISLA research of IDs in the instructed L2 learning of LCTLs.

Considering the low sample size, the results of the statistical analyses need to be interpreted with additional caution, particularly because of the pairwise deletion used throughout the analyses to maximize the number of observations. Researchers have pointed out that this method of dealing with missing data may lead to biased parameter estimates (Lodder, 2013). To address this issue, researchers have used more conservative methods such as listwise deletion

with larger samples (see, e.g., Kisselev et al., forthcoming). It is also important to note that even though the summer program had successfully used the ACTFL OPIs/OPIcs and WPTs and their internal lexicogrammar quizzes throughout its history, an investigation of these tests' validity would be beneficial and is considered a possible future direction for this project. In terms of the instruments used in this study, there were also several instances of potentially misfitting items and reliability issues with regards to the additional L2 exposure log (for more information, see Chapter 5). These issues are considered a limitation of the study. In a future replication study, researchers may consider refining the additional L2 exposure log and potentially collect this type of data from a series of interviews.

Another important future direction for this project would be an investigation of the relationships between L2 learners' IDs (see Serafini, 2017). The graduates of the summer program generally talked about IDs as a complex system and that their IDs were often interconnected. Several participants mentioned that their strengths in one area often helped them compensate for weaknesses in other areas, remaining consistent with the DST (e.g., de Bot et al., 2007; Serafini, 2017) and multidimensional perspectives on IDs (e.g., Sun & Zhang, 2020). For example, one of the participants mentioned that their higher motivation resulted in higher additional L2 exposure. Another participant mentioned that having a good memory but lacking intrinsic motivation resulted in lower additional L2 exposure for them. The interviewees often talked about motivation and L2 exposure as two interrelated categories. Even though it is impossible to study everything at once (Larsen-Freeman, 1997, 2012, 2014), future research may consider adding such important IDs as learning grit (Sudina & Plonsky, 2021); social dimensions (Duff, 2017); psychological dimensions and foreign language anxiety (Dewaele, 2017); L2 instructor individual characteristics (Gurzynski-Weiss, 2017); heritage learner status (Montrul &

Bowles, 2017), and the role of L2 aptitude (Li, 2017). Finally, it may also be important to investigate instructed L2 learning outside "WEIRD" communities. The Russian summer immersion program is considered an elite school that draws from universities all over the United States and abroad, or a community of students would generally fall under the category of "WEIRD" privileged learners (see Henrich et al, 2010). In SLA literature, such communities have been often referred to as "convenient" samples (e.g., Adringa & Godfroid, 2020), even though collecting data in this program was very challenging. Studies outside "WEIRD" communities would help promote the generalizability of ISLA research to larger populations and thus can further promote the research-practice dial (see, e.g., Sato et., 2021).

Conclusion

This study supported the application of the complexity/DST framework (a relatively new theory of SLA and ISLA) in ID research; illustrated the benefits of using mixed-methods research and multiple data points in ID studies of LCTLs such as Russian (previous research has often been quantitative and included only one data point); as well as conducting longitudinal ISLA studies in ecologically valid settings at different proficiency levels. Morphologically rich languages such as Russian require long learning periods for most native speakers of English (Looney & Lusin, 2019) and thus a longitudinal perspective is vital in this type of research. The results of this study supported the view of L2 learners' motivation (as measured by the L2MSS and General Motivation surveys) and L2 exposure (specifically exposure outside formal classes) as constructs showing both stable and dynamic patterns (referred to as attractor states in DST literature) and that different constituents of WM, motivation, L2 Russian exposure, and learning gains were engaged in complex relationships that varied at different curricular levels, language skills, and phases of the study. This dissertation includes a discussion of theoretical implications

with regards to DST as a future direction for ISLA research of IDs, as well as the discussion of the mixed-methods approach as a future direction for ISLA studies of LCTLs. This dissertation also includes a section on practical implications that discusses pedagogical aspects and implications for the development of background questionnaires and placement testing that may take L2 learners' IDs in cognitive abilities, affective factors, and L2 exposure into consideration.

APPENDICES

Appendix A: L2MSS survey

Please answer these questions about your motivation to study Russian.

	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9 (Strongly agree)
l can imagine myself as someone who is able to speak Russian.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l can imagine speaking Russian as if I were a native speaker.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I can imagine myself speaking Russian with international colleagues.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l can imagine myself living abroad and having a discussion in Russian.	\bigcirc	\bigcirc	0	\bigcirc	\bigcirc	0	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I can imagine myself studying in a university where all my courses are taught in Russian.	\bigcirc	\bigcirc	0	\bigcirc	0	0	\bigcirc	0	\bigcirc	\bigcirc

Please answer these questions about your motivation to study Russian.

	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9 (Strongly agree)
I chose to learn Russian despite others encouraging me to study something different (another language or a different subject entirely).	0	0	0	0	0	0	0	0	0	0
I am studying Russian even though most of my friends and family members don't value foreign language learning.	\bigcirc	0	0	0	0	0	0	0	0	0
I want to prove others wrong by becoming good at Russian.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l am studying Russian because it is a challenge.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I would like to reach a high proficiency in Russian, despite others telling me that it will be difficult or impossible.	\bigcirc	\bigcirc	0	0	0	0	\bigcirc	\bigcirc	\bigcirc	\bigcirc

Please answer these questions about your motivation to study Russian.

	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9 (Strongly agree)
If I fail to learn Russian, I'll be letting other people down.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
Studying Russian is important to me in order to gain the approval of my peers/teachers/family/boss.	\bigcirc	0	0	0	0	0	0	0	0	0
I consider learning Russian important because the people I respect think that I should do it.	\bigcirc	\bigcirc	0	0	0	0	0	0	0	0
It will have a negative impact on my life if I don't learn Russian.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I study Russian because close friends/family members of mine think it is important.	\bigcirc	\bigcirc	\bigcirc	0	0	0	0	0	\bigcirc	0

Rate your Russian learning experiences at your institution(s) BEFORE(if you have notstudied Russian beforeplease skip this question)

	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9 (Strongly agree)
I like the atmosphere of my Russian classes.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
I find learning Russian really interesting.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l always look forward to Russian classes.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
l would like to have more Russian lessons at school.	0	\bigcirc								
l really enjoy learning Russian.	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	0	\bigcirc	\bigcirc	\bigcirc	\bigcirc

Appendix B: General Motivation survey

Rate your CURRENT motivation to learn Russian.

	0 (Very Iow)	1	2	3	4	5	6	7	8	9 (Very high)
My motivation to learn Russian in order to communicate with Russian speaking people is	0	\bigcirc	0	0	0	0	0	\bigcirc	\bigcirc	0
My desire to learn Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
My motivation to learn Russian for practical purposes (e.g., to get a job) is	0	0	0	\bigcirc	\bigcirc	0	0	\bigcirc	\bigcirc	0
My interest in Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc
My motivation to learn Russian is	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc	\bigcirc

Appendix C: Additional L2 exposure log

This week at (Monday, August 2 - Friday, August 6), how much time did you spend in the following activities OUTSIDE formal classes?

	Hours this week
Online communication through social media/email using Russian	
Conversations with instructors in Russian	
Conversations with tutors in Russian	
Conversations with peers in Russian	
Homework in the Russian program	
Language learning apps to study Russian	
Using Russian while playing sports or video games	
Reading Russian literature	
Reading Russian news/blogs online	
Watching TV/movies/streaming services in Russian	
Listening to music in Russian	
Using Russian while participating in thematic clubs	
Attending Russian guest lectures	

Appendix D: Interview questions

- 1. To what extent has your motivation to study Russian changed over the course of your studies at [name of the summer program]? If so, how?
- 2. Do you envision yourself as someone who will speak Russian in the future? Has this vision changed over the course of your studies at [name of the summer program]? If so, how?
- 3. Has anyone affected your decision to study Russian (for example, your family members or friends)? To what extent was this external factor important for you when you started the program? Is it an important factor for you now?
- 4. Has anyone encouraged you NOT to study Russian? If yes, is it still the case after you finished the [name of the summer program] program?
- 5. How would you describe your Russian learning experiences at [name of the summer program]? What about your previous institution (if applicable)?
- 6. During the study, you mentioned that you were engaged in learning Russian OUTSIDE your formal classes at [name of the summer program] (for example, using language learning apps, attending guest lectures, watching Russian movies, listening to Russian music)? What activities OUTSIDE your formal classes were particularly helpful for learning Russian for you and why?
- 7. During the first phase of the study, you completed two working memory tests. To what extent do you think your working memory (or the ability to temporarily store information in your brain) affected your studies at [name of the summer program]? To what extent do you think your working memory helped you learn or hinder you from learning Russian at [name of the summer program]?

8. In my dissertation, I also explore the relationships between three categories of students' individual differences: working memory, motivation, and language exposure (or the number of hours you spend learning Russian). Do you think that your strength in one of these areas helped you compensate for your weaknesses in others?

Appendix E: IRB approval

MICHIGAN STATE UNIVERSITY

ACKNOWLEDGEMENT OF A MODIFICATION FOR AN EXEMPT DETERMINATION **Revised Common Rule**

June 15, 2021

Shawn D Loewen To:

Re: MSU Study ID: 00004904 Principal Investigator: Shawn D Loewen Category: Exempt 2(ii) Submission: Modification 00003984 Limited IRB Review: Not Required.

Title: The Interplay of Aptitude, Working Memory, Motivation, and Exposure in Russian Second Language Acquisition: A Compensatory Approach

In general, investigators are not required to submit changes to the IRB office once a research study is designated as exempt as long as those changes do not affect the exempt category or criteria for exempt determination (changing from exempt status to expedited or full review, changing exempt category) or that may substantially change the focus of the research study such as a change in hypothesis or study design.



This letter acknowledges the change you provided for the above listed study. The change does not affect the exempt category, criteria for exempt determination, or substantially change the focus of the research study.

Office of Regulatory Affairs Human Research Protection Program

> 4000 Collins Road Suite 136 Lansing, MI 48910

517-355-2180 Fax: 517-432-4503 Email: irb@msu.edu www.hrpp.msu.edu

participants' placement and exit tests in the summer program: FERPA regulations apply. Principal Investigator (PI) Responsibilities: The PI assumes the responsibilities

Additional subjects will participate in a short interview and researchers will access

for the protection of human subjects in this study as outlined in Human Research Protection Program (HRPP) Manual Section 8-1, Exemptions.

Continuing Review: Exempt studies do not need to be renewed.

Modifications: In general, investigators are not required to submit changes to the Michigan State University (MSU) Institutional Review Board (IRB) once a research study is designated as exempt as long as those changes do not affect the exempt category or criteria for exempt determination (changing from exempt status to expedited or full review, changing exempt category) or that may substantially change the focus of the research study such as a change in hypothesis or study design. See HRPP Manual Section 8-1, Exemptions, for examples. If the study is modified to add additional sites for the research, please note that you may not

REFERENCES

REFERENCES

- Ahmadian, M. J. (2020). Explicit and implicit instruction of refusal strategies: Does working memory capacity play a role? *Language Teaching Research*, 24(2), 163-188. https://doi.org/10.1177/1362168818783215
- Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, 40, 134-142. https://doi.org/10.1017/S0267190520000033
- Baddeley, A. (2010). Working memory. *Current biology*, 20(4), R136-R140. https://doi.org/10.1016/j.cub.2009.12.014
- Berghoff, R. (2022). L2 processing of filler-gap dependencies: Attenuated effects of naturalistic L2 exposure in a multilingual setting. *Second Language Research*, *38*(2), 373-393. https://doi.org/10.1177/0267658320945757
- Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00081
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005-2014: Understanding a publication surge and a changing landscape. *System*, 55, 145-157. https://doi.org/10.1016/j.system.2015.10.006
- Brooks, D. W., & Shell, D. F. (2006). Working memory, motivation, and teacher-initiated learning. *Journal of Science Education and Technology*, 15(1), 17-30. https://doi.org/10.1007/s10956-006-0353-0
- Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, 41(2), 281-296. https://doi.org/10.3758/s13421-012-0262-9
- Csizér, K. (2017). Motivation in the L2 classroom. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 418-432). New York, NY, USA: Routledge.
- de Bot, K. (2015). Rates of change: Timescales in second language development. In Z. Dörnyei,
 P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 29–37). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783092574-006
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, *10*(1), 7-21. https://doi.org/10.1017/S1366728906002732

- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2022). Formal versus informal L2 learning: How do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children? *Studies in Second Language Acquisition*, *44*(1), 87-111. https://doi.org/10.1017/S0272263121000097
- DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62*, 189-200. https://doi.org/10.1111/j.1467-9922.2012.00712.x
- Dengub, E., Dubinina, I., & Merrill, J. (Eds.) (2020). *The art of teaching Russian*. Washington, DC, USA: Georgetown University Press.
- Denhovska, N., & Serratrice, L. (2017). Incidental learning of gender agreement in L2. *Journal* of Psycholinguistic Research, 46(5), 1187-1211. https://doi.org/10.1007/s10936-017-9487-x
- Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning*, 66(1), 159-190. https://doi.org/10.1111/lang.12142
- Dewaele, J.-M. (2017). Psychological dimensions and foreign language anxiety. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 433-450). New York, NY, USA: Routledge.
- Di Santo, S., De Luca, V., Isaja, A., & Andreetta, S. (2020). Working memory training: Assessing the efficiency of mnemonic strategies. *Entropy*, 22(5), 577. https://doi.org/10.3390/e22050577
- Dörnyei, Z. (2009a). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691293-003
- Dörnyei, Z. (2009b). Individual differences: Interplay of learner characteristics and learning environment. *Language learning*, *59*(1), 230-248. https://doi.org/10.1111/j.1467-9922.2009.00542.x
- Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a Dynamic Systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). London, UK: Continuum. https://doi.org/10.5040/9781474212397.ch-009
- Dörnyei, Z. (2013). *The psychology of second language acquisition*. Oxford, UK: Oxford University Press.
- Dörnyei, Z. (2019). Towards a better understanding of the L2 learning experience, the Cinderella of the L2 Motivational Self System. *Studies in Second Language Learning and Teaching*, 9(1), 19-30. https://doi.org/10.14746/ssllt.2019.9.1.2

- Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self images, sensory styles, and imagery capacity across two target languages. *Language Learning*, 63(3), 437-462. https://doi.org/10.1111/lang.12005
- Dornyei, Z., & Kubanyiova, M. (2014). *Motivating learners, motivating teachers: Building vision in the language classroom*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781009024648
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York, NY, USA: Routledge/Taylor & Francis.
- Duff, P. A. (2017). Social dimensions and differences in instructed SLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 379-395). New York, NY, USA: Routledge.
- Duman, S. K., Yalçın, Ş., & Erçetin, G. (2021). Working memory and language aptitude in relation to listening strategy instruction in an instructed SLA context. *Annual Review of Applied Linguistics*, 1-10. https://doi.org/10.1017/s0267190521000040
- Eldh, A. C., Årestedt, L., & Berterö, C. (2020). Quotations in qualitative studies: Reflections on constituents, custom, and purpose. *International Journal of Qualitative Methods*, 19. https://doi.org/10.1177%2F1609406920969268
- Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language and Cognition*, *10*(1), 23-25. https://doi.org/10.1017/S1366728906002744
- Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, 34(1), 67-101. https://doi.org/10.1177%2F0267658316684903
- Fathi, J., Ahmadnejad, M., & Yousofi, N. (2019). Effects of blog-mediated writing instruction on L2 writing motivation, self-efficacy, and self-regulation: A mixed methods study. *Journal of Research in Applied Linguistics*, 10(2), 159-181.
- Foster, P., Bolibaugh, C., & Kotula, A. (2014). Knowledge of nativelike selections in a L2: The influence of exposure, memory, age of onset, and motivation in foreign language and immersion settings. *Studies in Second Language Acquisition*, 36(1), 101-132. https://doi.org/10.1017/s0272263113000624
- Gallo, F., Novitskiy, N., Myachykov, A., & Shtyrov, Y. (2021). Individual differences in bilingual experience modulate executive control network and performance: behavioral and structural neuroimaging evidence. *Bilingualism: Language and Cognition*, 24(2), 293-304. https://doi.org/10.2139/ssrn.3519852

- Gardner, R. C., Masgoret, A.-M., Tennant, J., & Mihic, L. (2004). Integrative motivation: Changes during a year-long intermediate-level language course. *Language Learning*, 54, 1-34. https://doi.org/10.1111/j.1467-9922.2004.00247.x
- Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA, USA: Newbury House.
- Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, *19*(1), 3-17. https://doi.org/10.1075/aila.19.03gas
- Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition*, *38*(2), 177-215. https://doi.org/10.1017/S0272263115000388
- Grey, S., Cox, J. G., Serafini, E. J., & Sanz, C. (2015). The role of individual differences in the study abroad context: Cognitive capacity and language development during short-term intensive language exposure. *The Modern Language Journal*, 99(1), 137-157. https://doi.org/10.1111/modl.12190
- Gurzynski-Weiss, L. (2017). L2 instructor individual characteristics. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 451-467). New York, NY, USA: Routledge.
- Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-Murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology* and Aging, 26(1), 92. https://doi.org/10.1037/a0021483
- Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing*, 40, 44-59. https://doi.org/10.1016/j.jslw.2018.03.001
- Hanzawa, K. (2021). Development of second language speech fluency in foreign language classrooms: A longitudinal study. *Language Teaching Research*, Article 13621688211008693. https://doi.org/10.1177/13621688211008693
- Hayashi, Y. (2019). Investigating effects of working memory training on foreign language development. *The Modern Language Journal*, 103(3), 665-685. https://doi.org/10.1111/modl.12584
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and brain sciences, *33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X
- Henry, A. (2017). L2 motivation and multilingual identities. *The Modern Language Journal*, 101(3), 548-565. https://doi.org/10.1111/modl.12412

- Hiver, P. (2015). Attractor states. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 20–28). Bristol, UK: Multilingual Matters.
- Ibarra Santacruz, D., & Martínez Ortega, D. (2018). Can working memory strategies enhance English vocabulary learning? *How*, 25(2), 29-47. https://doi.org/10.19183/how.25.2.410
- Indrarathne, B., & Kormos, J. (2018). The role of working memory in processing L2 input: Insights from eye-tracking. *Bilingualism: Language and Cognition*, 21(2), 355-374. https://doi.org/10.1017/S1366728917000098
- Isabelli-García, C., & Lacorte, M. (2016). Language learners' characteristics, target language use, and linguistic development in a domestic immersion context. *Foreign Language Annals*, 49(3), 544-556. https://doi.org/10.1111/flan.12215
- Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, 18(1), 3-20. https://doi.org/10.1177/1525822x05282260
- Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching*, *10*(1), 89-109. https://doi.org/10.14746/ssllt.2020.10.1.5
- Kim, J., Zhao, H., & Diskin-Holdaway, C. (2022). Willingness to communicate and second language fluency: Korean-speaking short-term sojourners in Australia. *Languages*, 7(2), 112. https://doi.org/10.3390/languages7020112
- Kisselev, O., Dubinina, I., & Polinsky, M. (2020). Form-focused instruction in the heritage language classroom: Toward research-informed heritage language pedagogy. *Frontiers in Education*, *5*, Article 53. https://doi.org/10.3389/feduc.2020.00053
- Kisselev, O., Soyan, R., Pastushenkov, D., & Merrill, J. (forthcoming). Measuring writing development and proficiency gains using indices of lexical and syntactic complexity: Evidence from longitudinal Russian learner corpus data.
- Klimanova, L., & Dembovskaya, S. (2013). L2 identity, discourse, and social networking in Russian. *Language Learning & Technology*, *17*(1), 69-88. http://dx.doi.org/10125/24510
- LaFlair, G. T., Isbell, D., May, L. D. N., Gutierrez Arvizu, M. N., & Jamieson, J. (2017). Equating in small-scale language testing programs. *Language Testing*, 34(1), 127-144. https://doi.org/10.1177%2F0265532215620825
- Lai, H. Y. T. (2013). The motivation of learners of English as a Foreign Language revisited. *International Education Studies*, 6(10), 90-101. https://doi.org/10.5539/ies.v6n10p90

- Lanvers, U. (2016). Lots of selves, some rebellious: Developing the self discrepancy model for language learners. *System*, *60*, 79-92. https://doi.org/10.1016/j.system.2016.05.012
- Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? *Language Teaching*, 45, 202–214. https://doi.org/10.1017/s0261444811000061
- Larsen-Freeman, D. (2014). Ten 'lessons' from Complex Dynamic Systems Theory: What is on offer. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 11–19). Bristol, UK: Multilingual Matters. http://dx.doi.org/10.21832/9781783092574-004
- Lee, M., & Revesz, A. (2021). The role of working memory in attentional allocation and grammatical development under textually-enhanced, unenhanced and no captioning conditions. *Journal for the Psychology of Language Learning*, 3(1), 6-25. https://doi.org/10.52598/jpll/3/1/1
- Leona, N. L., van Koert, M. J., van der Molen, M. W., Rispens, J. E., Tijms, J., & Snellings, P. (2021). Explaining individual differences in young English language learners' vocabulary knowledge: The role of Extramural English Exposure and motivation. *System*, 96, Article 102402. https://doi.org/10.1016/j.system.2020.102402
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on preregistration. *International Review of Social Psychology*, 32(1), Article 5. https://psycnet.apa.org/doi/10.5334/irsp.289
- Li, S. (2017). Cognitive differences and ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge* handbook of instructed second language acquisition (pp. 396-417). New York, NY, USA: Routledge.
- Li, S., Hiver, P., & Papi, M. (Eds.). (2022). The Routledge handbook of second language acquisition and individual differences. Taylor & Francis. https://doi.org/10.4324/9781003270546-2
- Li, J. T., & Tong, F. (2019). Multimedia-assisted self-learning materials: The benefits of Eflashcards for vocabulary learning in Chinese as a foreign language. *Reading and Writing*, 32(5), 1175-1195. https://doi.org/10.1007/s11145-018-9906-x
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861-883. https://doi.org/10.3758/s13423-013-0565-2
- Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom? SAGE Open, 5(4), 1-11. https://doi.org/10.1177%2F2158244015607352

- Lodder, P. (2013). To impute or not impute, that's the question. In J. G. Mellenbergh, and H. J. Adèr (Eds.). *Advising on research methods: Selected topics*. Huizen, The Netherlands: Johannes van Kessel Publishing.
- Loewen, S., & Hui, B. (2021). Small samples in instructed second language acquisition research. *The Modern Language Journal*, *105*(1), 187-193. https://doi.org/10.1111/modl.12700
- Loewen, S., & Sato, M. (Eds.). (2017). *The Routledge handbook of instructed second language acquisition*. New York, NY, USA: Routledge. https://doi.org/10.4324/9781315676968
- Looney, D., & Lusin, N. (2019). Enrollments in languages other than English in United States institutions of higher education, summer 2016 and fall 2016: Final report. Modern Language Association of America.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60(3), 501-533. https://doi.org/10.1111/j.1467-9922.2010.00565.x
- Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, 62(3), 704-740. https://doi.org/10.1111/j.1467-9922.2011.00649.x
- Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, 40(3), 651-675. https://doi.org/10.1017/s0272263117000341
- Marsden, E., Mackey A., & Plonsky, L. (2016). The IRIS Repository: Advancing research practice and methodology. In A. Mackey & E. Marsden (Eds.), Advancing methodology and practice: The IRIS Repository of Instruments for Research into Second Languages (pp. 1-21). New York, NY, USA: Routledge.
- Martinsen, R. A., Baker, W., Bown, J., & Johnson, C. (2011). The benefits of living in foreign language housing: The effect of language use and second-language type on oral proficiency gains. *The Modern Language Journal*, 95(2), 274-290. https://doi.org/10.1111/j.1540-4781.2011.01180.x
- Masgoret, A.-M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53, 123-63. https://doi.org/10.1111/1467-9922.00227
- Merrill, J. (2020). Student perceptions of a total immersion environment. In J. Bown, W. Smemoe, & D. P. Dewey, (Eds.), *Language learning in foreign language houses: Immersion and integration* (pp. 183-217). Auburn, AL, USA: International Association for Language Learning Technology.

- Merrill, J., Dengub, E., & Pastushenkov, D. (2021). Language gains in intensive synchronous online and face-to-face Russian immersion programs: A Comparison. *Russian Language Journal*, *71*(2), 1-34. https://scholarsarchive.byu.edu/rlj/vol71/iss2/6
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.
- Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 29(4), 578-596. https://doi.org/10.1093/applin/amm056
- Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, *35*(4), 463-482. https://doi.org/10.1093/applin/amu024
- Montrul, S., & Bowles, M. (2017). Instructed heritage language acquisition. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 488-502). New York, NY, USA: Routledge.
- Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: Investigating motivation as a time-varying predictor of pronunciation development. *The Modern Language Journal*, 102(1), 199-217. https://doi.org/10.1111/modl.12461
- Nielson, K. B., & DeKeyser, R. (2019). Working memory and planning time as predictors of fluency and accuracy. *Journal of Second Language Studies*, 2(2), 281-316. https://doi.org/10.1075/jsls.19004.bro
- Odom, L. R., & Morrow, Jr, J. R. (2006). What's this R? A correlational approach to explaining validity, reliability and Objectivity coefficients. *Measurement in Physical Education and Exercise Science*, *10*(2), 137-145. https://doi.org/10.1207/s15327841mpee1002_5
- Oliver, R., Nguyen, B., & Sato, M. (2017). L2 instructor individual characteristics. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 468-487). New York, NY, USA: Routledge.
- Papi, M., & Hiver, P. (2020). Language learning motivation as a complex dynamic system: A global perspective of truth, control, and value. *The Modern Language Journal*, 104(1), 209-232. https://doi.org/10.1111/modl.12624
- Pastushenkov, D. (2020). Second language acquisition in the digital world: Language app use as an individual difference. *Journal of Global Literacies, Technologies, and Emerging Pedagogies*. 6(1), 1103-1119.

- Pastushenkov, D., & McIntyre, T. (2020). Life after language immersion: Two very different stories. *The Journal of the National Council of Less Commonly Taught Languages*. 27, 1-32.
- Pastushenkov, D., Camp, C., Zhuchenko, I., & Pavlenko, O. (2021). Shared and different L1 background, L1 use, and peer familiarity as factors in ESL pair interaction. *TESOL Journal*, 12(2), e538. https://doi.org/10.1002/tesj.538
- Pattemore, A., & Muñoz, C. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System*, *93*, Article 102303. https://doi.org/10.1016/j.system.2020.102303
- Pawlak, M. (2012). Individual differences in language learning and teaching: Achievements, prospects and challenges. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 19-46). Heidelberg, Germany: Springer.
- Pfenninger, S. E., & Singleton, D. (2016). Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research*, *32*(3), 311-345. https://doi.org/10.1177/0267658315624476
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. Language Learning, 64(4), 878-912. https://doi.org/10.1111/lang.12079
- Ranta, L., & Meckelborg, A. (2013). How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *Canadian Modern Language Review*, 69(1), 1-33. https://doi.org/10.3138/cmlr.987
- Rifkin, B. (2005). A ceiling effect in traditional classroom foreign language instruction: Data from Russian. *The Modern Language Journal*, *89*(1), 3-18. https://doi.org/10.1111/j.0026-7902.2005.00262.x
- Ruiz, S., Rebuschat, P., & Meurers, D. (2021). The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. *Language Teaching Research*, 25(4), 510-539. https://doi.org/10.1177/1362168819872859
- Sagarra, N. (2019). Individual differences in working memory and instructed SLA. In R. P. Leow (Ed.), *The Routledge Handbook of Second Language Research in Classroom Learning* (pp. 452-463). New York, NY, USA, Routledge.
- Sainani, K. L. (2015). Dealing with missing data. *PM&R*, 7(9), 990-994. https://doi.org/10.1016/j.pmrj.2015.07.011
- Saito, K., & Hanzawa, K. (2018). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*, 22(4), 398-417. https://doi.org/10.1177/1362168816679030

- Sato, M., & Csizér, K. (2021). Introduction: Combining learner psychology and ISLA research: Intersections in the classroom. *Language Teaching Research*, 25(6), 839-855. https://doi.org/10.1177%2F13621688211044237
- Sato, M., & Lara, P. (2019). Interaction vision intervention to increase second language motivation: A classroom study. In M. Sato & S. Loewen (Eds.), *Evidence-based second language pedagogy* (pp. 287-313). New York, NY, USA: Routledge. https://doi.org/10.4324/9781351190558-13
- Sato, M., & Loewen, S. (2019a). Toward evidence-based second language pedagogy: Research proposals and pedagogical recommendations. In M. Sato & S. Loewen (Eds.), *Evidence-Based Second Language Pedagogy* (pp. 1-23). New York, NY, USA: Routledge.
- Sato, M., & Loewen, S. (2019b). Methodological strengths, challenges, and joys of classroombased quasi-experimental research: Metacognitive instruction and corrective feedback. In R. DeKeyser & G. Prieto Botana (Eds.), *Doing SLA research with implications for the classroom: Reconciling methodological demands and pedagogical applicability* (pp. 31– 54). John Benjamins. https://doi.org/10.1075/Illt.52.03sat
- Sato, M., Loewen, S., & Pastushenkov, D. (2021). 'Who is my research for?': Researcher perceptions of the research-practice relationship. *Applied Linguistics*, 1-26. https://doi.org/10.1093/applin/amab079
- Sagarra, N. (2017). Longitudinal effects of working memory on L2 grammar and reading abilities. *Second Language Research*, *33*(3), 341-363. https://doi.org/10.1177/0267658317690577
- Segalowitz, N., & Trofimovich, P. (2012). Second language processing. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 179–192). New York, NY, USA: Routledge/Taylor& Francis.
- Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *The Modern Language Journal*, 101(2), 369-390. https://doi.org/10.1177%2F0267658317690577
- Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, 38(4), 607-646. https://doi.org/10.1017/S0272263115000327
- Schraw, G., Brooks, D. W., & Crippen, K. J. (2005). Using an interactive, compensatory model of learning to improve chemistry teaching. *Journal of Chemical Education*, 82(4), 637. https://doi.org/10.1021/ed082p637

- Sparks, R. L., Patton, J., & Ganschow, L. (2012). Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences*, 22(4), 463-472. https://doi.org/10.1016/j.lindif.2012.03.009
- Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31, 532–553. https://doi.org/10.1093/applin/amq001
- Stanovich, K. E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education*, 5(3), 11-19.
- Sudina, E., & Plonsky, L. (2021). Language learning grit, achievement, and anxiety among L2 and L3 learners in Russia. *ITL-International Journal of Applied Linguistics*, 172(2), 161-198. https://doi.org/10.1075/itl.20001.sud
- Sun, Y. (2019). An analysis on the factors affecting second language acquisition and its implications for teaching and learning. *Journal of Language Teaching and Research*, 10(5), 1018-1022. http://dx.doi.org/10.17507/jltr.1005.14
- Sun, P. P., & Zhang, L. J. (2020). A multidimensional perspective on individual differences in multilingual learners' L2 Chinese speech production. *Frontiers in Psychology*, 11(59), 1-14. https://doi.org/10.3389/fpsyg.2020.00059
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2
- Taguchi, N. (2018). Description and explanation of pragmatic development: Quantitative, qualitative, and mixed methods research. *System*, 75, 23-32.
- Thompson, A. S. (2017). Language learning motivation in the United States: An examination of language choice and multilingualism. *The Modern Language Journal*, *101*(3), 483-500. https://doi.org/10.1111/modl.12409
- Thompson, A. S. (2021). Conceptualizing the anti-ought-to self: Background and new directions. *Revue TDFLE*, (78). https://doi.org/10.34745/numerev_1699
- Thompson, A. S., & Vásquez, C. (2015). Exploring motivational profiles through language learning narratives. *The Modern Language Journal*, 99(1), 158-174. https://doi.org/10.1111/modl.12187
- Tigchelaar, M., Bowles, R. P., Winke, P., & Gass, S. (2017). Assessing the validity of ACTFL can-do statements for spoken proficiency: A Rasch analysis. *Foreign Language Annals*, *50*(3), 584-600. https://doi.org/10.1111/flan.12286

- Toth, P., & Davin, K. J. (2016). The sociocognitive imperative of L2 pedagogy. *Modern Language Journal*, *100* (Supplement 2016), 148-168. https://doi.org/10.1111/modl.12306
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28(1), 1-30. https://doi.org/10.1017/s0272263106060013
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? Journal of Memory and Language, 28(2), 127-154.
- Ushioda, E., & Dörnyei, Z. (2012). Motivation. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 396–409). New York, NY, USA: Routledge/Taylor & Francis.
- van den Noort, M. W., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist*, 11(4), 289.
- van Geert, P. L. C. (1994). *Dynamic systems of development: Change between complexity and chaos*. New York, NY, USA: Harvester Wheatsheaf.
- van Geert, P. L. C. (1995). Growth dynamics in development. In R. F. Port & T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 313–338). Cambridge, MA: Bradford Books/The MIT Press.
- Verhaeghen, P., Martin, M., & Sędek, G. (2012). Reconnecting cognition in the lab and cognition in real life: The role of compensatory social and motivational factors in explaining how cognition ages in the wild. *Aging, Neuropsychology, and Cognition, 19*(1-2), 1-12. https://doi.org/10.1080/13825585.2011.645009
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA, USA: Harvard University Press.
- Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12-23. https://doi.org/10.1016/j.learninstruc.2013.04.002
- Waninge, F., Dörnyei, Z., & de Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *Modern Language Journal*, 98, 704-723. https://doi.org/10.1111/modl.12118
- Wiley, J., Sanchez, C. A., & Jaeger, A. J. (2014). The individual differences in working memory capacity principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook* of multimedia learning (2nd ed., pp. 598-619). New York, NY, USA: Cambridge University Press.

- Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal*, *97*(1), 109-130. https://doi.org/10.1111/j.1540-4781.2013.01428.x
- Zaykovskaya, I., Rawal, H., & De Costa, P. I. (2017). Learner beliefs for successful study abroad experience: A case study. *System*, *71*, 113-121. https://doi.org/10.1016/j.system.2017.09.020
- Zhu, Y., Fung, A. S., & Wang, H. (2012). Memorization effects of pronunciation and stroke order animation in digital flashcards. *CALICO Journal*, 29(3), 563-577. https://doi.org/10.11139/cj.29.3.563-577