Improving the Reading Achievement of Language Minority and Disadvantaged Youth At Risk of

Academic Failure

by

Linda Iwenofu

thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy Graduate Department of Applied Psychology and Human Development Ontario Institute for Studies in Education University of Toronto

© Copyright by Linda Iwenofu 2019

Improving the Reading Achievement of Language Minority and Disadvantaged Youth At Risk of Academic Failure

Linda Iwenofu

Doctor of Philosophy

Department of Applied Psychology and Human Development

University of Toronto

2019

Abstract

BACKGROUND: Variability in individuals' response to intervention can contribute to smaller intervention effects. Integration of quantitative and qualitative findings can be instrumental in elucidating person-level and broader contextual issues related to differential intervention efficacy and inform overall intervention utility. In order to assess ecological factors implicated in differential response to intervention, three studies were conducted that together comprised a comprehensive program evaluation of the Vocabulary Learning Project, a manualized reading intervention targeting the vocabulary and reading comprehension skills of academically at-risk language-minority and economically disadvantaged high-school aged youth.

METHODS: A sequential embedded quasi-experimental mixed methods research design consisting of three distinct yet interrelated phases was used. Across the three research phases, mixed analyses of variance, hierarchical linear regression and multi-case study analyses were conducted to assess intervention effects, factors predictive of outcome gains and contextual factors differentiating outcomes, respectively. RESULTS: Findings from the first two studies indicated that the intervention was differentially effective based on multiple factors at various contextual levels. At the person-level context, pre-intervention language comprehension skills, pre-intervention motivation to read, achievement orientation, academic self-concept and sense of future aspirations were implicated factors. At the intervention program level, investment in positive program outcomes, tutoring group climate and participant resourcefulness were identified factors. Within the peer and family social context, reliance on peers and nature of parental support were differentiating factors. At the broader school, community and cultural context, school perceptions, school and community engagement, as well as youths' sense of cultural identity were factors that differentiated program effects. Findings from the third, integrative study resulted in the identification of a unifying meta-theme of motivation as a key factor underlying the differential responses to the VLP intervention.

CONCLUSIONS: Consideration of the multiple contexts navigated by culturally and linguistically diverse youth is critical for maximizing intervention effects.

IMPACT: This study provides multiple insights about opportunities worthy of consideration in optimally designing instructional interventions targeting marginalized youth populations in the Canadian context, and suggests a useful methodological approach for evaluating such programs.

iii

Acknowledgements

There are several people without whom this research endeavour would have been impossible. Dr. Esther Geva was not only my thesis advisor but also a significant mentor throughout my graduate studies. I appreciate her immensely for teaching and modelling that most things can be achieved with persistence, determination and hard work. As a result, I feel ever confident that the sky is the limit for me. Thank you for helping me navigate what was a complex yet rewarding research project, and for guiding me through my clinical program. You genuinely continue to inspire me.

My sincere gratitude goes to the rest of my supervisory committee. Dr. Judy Wiener, I thank you for betting on me right from the beginning, for your guidance through all phases of this work and for being an esteemed mentor who played a significant role in my professional development. Dr. Alexandra Gottardo, I sincerely appreciated your tremendous expertise and invaluable feedback throughout all phases of this project, from its conception onward.

I must also thank the full team of Co-Investigators who conceptualized and designed the VLP: Dr. Esther Geva, Dr. Alexandra Gottardo, Dr. Fataneh Farnia and Dr. Xi Becky Chen. Additionally, many thanks to the core team of graduate coordinators who gave daily life and sustenance to this project, often giving up so much of themselves and their time for its sake: Adrian Pasquarella, Christie Fraser, Mahshid Azimi and Tamara Kornacki.

Deserving of significant praise are the numerous and invaluable volunteers who comprised the extended project team, including volunteer research assistants and program tutors who tirelessly worked to do their part in executing and managing this project, and who supported the youth participants on an ongoing basis. To the student VLP participants, their parents and Pathways to Education staff – this project would not have been possible without the gracious commitment of your time and expertise.

iv

Finally, to Chris and our boys for being my foundation over the course of graduate school – thank you for your unwavering love, faith and support.

Acknowlee	lgementsiv
Table of C	ontentsvi
List of Tal	lesix
List of Fig	ıresxi
List of Ap	oendicesxii
Chapter 1	General Introduction1
1.1 Sta	tement of the Problem
1.2 Stu	dy Rationale
1.3 Ov	erview of Studies
1.4 Res	earch Questions
1.5 Mi	xed Methods Approach
1.6 Ov	erview of Chapters
Chapter 2	Reading Comprehension Development in Academically At-Risk Adolescents:
Chapter 2 The Role of	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1)13
Chapter 2: The Role of 2.1 Inte	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) oduction
Chapter 2: The Role of 2.1 Intr 2.1.1	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) oduction 13 Demographic Context. 13
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) oduction 13 Demographic Context 13 Theoretical Context 16
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3 2.1.4	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28 Research Questions 29
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3 2.1.4 2.2 Meth	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28 Research Questions 29 ods 30
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3 2.1.4 2.2 Meth 2.2.1	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28 Research Questions 29 nods 30 Community Context 30
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3 2.1.4 2.2 Meth 2.2.1 2.2.2	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) oduction 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28 Research Questions 29 nods 30 Community Context 30 The 'Vocabulary Learning Project' Intervention 32
Chapter 2: The Role of 2.1 Intr 2.1.1 2.1.2 2.1.3 2.1.4 2.2 Meth 2.2.1 2.2.2 2.2.3	Reading Comprehension Development in Academically At-Risk Adolescents: f Intervention and Individual Differences (Study 1) 13 oduction 13 Demographic Context 13 Theoretical Context 16 Summary and Rationale for Present Study 28 Research Questions 29 nods 30 Community Context 30 The 'Vocabulary Learning Project' Intervention 32 Research Design 37

Table of Contents

2.2.5	Measures	46
2.2.6	Preliminary data analysis	51
2.2.7	Program Fidelity	62
2.2.8	Data Analysis	63
2.3 Re	sults	68
2.3.1	Exploration of the effects of a reading intervention on youth vocabulary and reading	
com	prehension	. 68
2.3.2	Evaluation of intervention effects for subgroups of participants	72
2.3.3	Examination of predictors of reading comprehension gains over time	79
2.4 D	iscussion	86
Chapter	3: An Exploration of Contextual Factors Associated with Response to Reading	
Interven	tion with Marginalized Adolescents (Study 2)	97
3.1 II	ntroduction	97
3.1.2	Research Objectives	105
3.2 N	lethod	107
3.2.1	Methodological Approach: Multiple Case Analysis	107
3.2.2	Positioning of the Primary Researcher	109
3.2.3	Community Context	111
3.2.4	VLP Intervention	112
3.2.5	Procedure	112
3.2.6	Measures	114
3.2.7	Participants	116
3.2	.7.1 Procedure for purposive selection of cases	119
3.2.8	Data Analysis	125
3.2.9	Assessing Quality of Findings	131

3.3 Findings	133
3.4 Summary and Interpretation of Findings	155
Chapter 4 – A Mixed Methods Evaluation of the Vocabulary Learning Project (VLP):	
Integration of Findings from Studies 1 and 2 (Study 3)	165
4.1 Introduction	165
4.1.1 Purpose of the current study	166
4.1.2 Research Question	166
4.2 Methods	167
4.2.1 Research Design: Mixed Methods Approach	167
4.2.2 Data Analysis	170
4.2.3 Measures	171
4.3 Results	172
4.4 Discussion	178
4.4.1 Comparison of findings across Studies 1 and 2	178
4.4.2 General Discussion	181
References	188
Appendices	208

List of Tables

Table 1 VLP Instructional Strategies and Activities
Table 2 VLP Session Overview
Table 3 Characteristics of Study Participants and Dropouts 52
Table 4 Characteristics of Intervention Completers and Non-Completers 53
Table 5 Characteristics of Intervention Cohorts (Cohorts 1, 2 and 3)
Table 6 Characteristics of Control Cohorts (Cohorts 2, 3 and 4) 59
Table 7 Baseline Characteristics of Entire Study Sample, by Study Condition60
Table 8 Overview of Research Objectives and Data Analyses 67
Table 9 Means, Standard Deviations, and Standard Scores of Outcome Measures 68
Table 10Means, Standard Deviations and Results of Mixed ANOVAs for Intervention Effects 69
Table 11 Means, Standard Deviations and Results of Mixed ANOVAs for Selected Intervention
and Control Participant Subgroups at Pre- and Post-test
Table 12 Factor Analysis of Theoretically Correlated Predictor Variables (n=79)
Table 13 Descriptive Characteristics of Predictors of Reading Comprehension Gains
Table 14 Correlations of Demographic (age, years in Canada), Cognitive (nonverbal ability),
Linguistic (word reading, language comprehension) and Psychological (motivation) variables
with Reading Comprehension Residual Gain Scores (n=79)

Table 15	Hierarchical Linear Regression Predicting Reading Comprehension Gains on the	
WRMT-R	Passage Comprehension Task, After Controlling Demographic and Cognitive Factors	
(n=79)		. 85
Table 16	Characteristics of Interview Respondents and Non-Respondents from Cohort 2	. 117
Table 17	Characteristics of Selected Cases and Non-Selected Cases	. 118
Table 18	Characteristics of Improvers and Non-Improvers	. 125
Table 19	Within- and Across-Case Analytic Strategies	. 127
Table 20	Assessment of Research Quality	. 132
Table 21	Summary of Themes	. 134
Table 22	Overall Mixed Method Study Summary	. 169
Table 23	Measures Used for Integration of Findings	. 172
Table 24	Methods Matrix: Profile of Case Characteristics for Mixed Method Integration	. 173

List of Figures

Figure 1	MMR Model: Sequential Embedded Quasi-Experimental Design	11
Figure 2	An expanded framework for understanding reading comprehension in language-	
minority	learners	24
Figure 3	Flow Diagram of Participant Recruitment	41
Figure 4	Case selection process	124

List of Appendices

Appendix A	Comparison of Program Implementation Across Intervention Cohorts	208
Appendix B	Measures: Family Background Questionnaire, Pre-Interview Questionnaire and	
Semi-structu	red Student Interview Guide	211
B.1	Family Background Questionnaire	211
B.2	Pre-Interview Questionnaire	215
B.3	Semi-structured Student Interview Guide	217

Chapter 1: General Introduction

1.1 Statement of the Problem

Youth living in urban marginalized, economically underprivileged communities are at risk for academic underachievement and consequent high school dropout (Tilleczek, 2008, 2013). In fact, there is a widely acknowledged and persistent achievement gap between students attending schools within impoverished communities, and those who do not (Dei, 2008; Toronto District School Board (TDSB), 2008; Tilleczek, 2008, 2013). Poor urban communities also tend to house disproportionate numbers of immigrants from a variety of ethnic backgrounds, whose home language is not English (Statistics Canada, 2013). For these culturally and linguistically diverse immigrant youth, the risk of academic underachievement is further compounded by their language minority status (August & Shanahan, 2006; Cummins, 1991; Cummins & Persad, 2014). A complex interaction of various factors have been widely acknowledged to contribute to the "at-risk" status of racialized and immigrant youth within underprivileged communities, including such issues as racialized and gendered poverty, discrimination, low levels of motivation and engagement with the mainstream school curriculum, impoverished school environments, and high teacher turnover (Dei, 2008; Tilleczek, 2008, 2013). As a result of some of these problems, the high school dropout rate among this population of students is alarmingly high. To illustrate, up to 40% of students who self-identify as Black within the TDSB, the largest school board in Canada, drop out of high school (TDSB, 2008; Linley, 2004).

It should be noted that in the context of this study, the term "at-risk" is used to refer to youth who are at risk of dropping out of high school for a variety of reasons, including having low academic achievement, recent immigrant status, language and/or visible minority status, experiencing disadvantages such as poverty, marginalization, and having to attend impoverished schools within underprivileged neighborhoods. Thus to be "at-risk", as described here, is to be generally vulnerable to a myriad of disadvantages. In turn, these disadvantages may either increase a youth's susceptibility to academic failure and/or high school dropout, or in other cases trigger resilience and academic success in the face of adversity.

A particular aspect of the academic experience of at-risk youth that is of relevance to the current study is reading achievement. The reading skill development of ethnic and linguistic minority youth is an area of growing concern to researchers, especially the development of those advanced skills most important for navigating academic texts at the high school level (i.e., reading comprehension). The gap in reading achievement of such students compared to students residing in less disadvantaged neighborhoods or who have resided in Canada for a longer time (including Canadian-born, native English speaking youth) has been well documented (August & Shanahan, 2006; Human Resources Skills and Development Canada [HRSDC], 2008; Organization for Economic Co-operation and Development [OECD], 2018a; 2018b). As such, there is growing concern among researchers and educators about how to effectively close the literacy gap among vulnerable youth. Over the last few years, intervention efforts have been gradually building in order to target this reading achievement gap and help decrease the likelihood of dropout and consequent poor life outcomes (Snow & Biancarosa, 2003; Tilleczek, 2013; Lesaux, Kieffer, Kelley & Harris, 2014). There is a substantial body of evidence indicating that high school dropout is associated with unfavorable consequences for the individual as well as society. These include heightened risks of unemployment, increased health problems, higher rates of criminality and incarceration (De Witte & Rogge, 2013; Solga, 2002; Lochner & Moretti, 2004). These outcomes in turn represent a significant cost to society and contribute to overall declines in economic growth (De Witte & Rogge, 2013).

1.2 Study Rationale

Intervention efforts that target the development of specific reading skills theorized to be implicated in the comprehension of academic texts among at-risk adolescent learners are steadily increasing (e.g., Carlo et al., 2004; Dalton, Proctor, Uccelli, Mo, & Snow, 2011; Lesaux, Kieffer, Faller, & Kelley, 2010; Lesaux et al., 2014; Lubliner & Smetana, 2005; Snow, Lawrence, & White, 2009; Townsend & Collins, 2009; Cumming, 2012; Hwang, Lawrence, Mo & Snow, 2015). A growing body of research suggests that a focus on the development of academic vocabulary and reading comprehension strategies may be a worthwhile instructional target for literacy interventions. However, to date the existing research base is insufficient for making conclusions about the utility of targeting these areas (Lesaux et al., 2014). Historically, reading research and instructional initiatives have focused on young children; however, growing concerns about the need to develop evidence-based approaches to promoting the literacy skill development of adolescent students have prompted the development of new initiatives targeting this vulnerable population of students. As intervention efforts grow, there is in turn a growing need to identify the appropriate components of reading development to target in interventions, in order to produce the greatest gains in students - particularly those who are not only culturally and linguistically diverse but also marginalized and economically underprivileged. In addition to identifying the appropriate instructional components of reading to target in intervention, and evaluating interventions' overall impact on reading outcomes, the current study poses the argument that it is equally important to delineate the individual and contextual factors that influence responsiveness to intervention, in order to best tailor interventions to students' needs. Typically, even when the overall population level efficacy of an intervention is examined, little is known about the contextual processes that influence program implementation, as well as the factors that help explain how and why the intervention impacted students differentially (Lesaux

et al., 2014). There is a need for large scale evaluation research that attends to this, particularly when the adolescents of interest are those whose learning needs have often historically been unmet by the broader mainstream curriculum for one reason or another (see Dei, 2008; Tilleczek, 2013 for examples).

Only a few of the existing reading intervention evaluation studies are experimental in nature (e.g., August, Branum-Martin, Cardenas-Hagan & Francis, 2009; Lesaux et al., 2014; Townsend & Collins, 2009; Vaughn et al., 2009). The limitation of most of these studies is that they solely espouse quantitative methodology and have not been conducted with adolescents in the secondary school grades, especially within the Canadian context. There may be benefit in evaluating the differential outcomes of reading interventions using a mixed methods approach. Mixed methodology allows researchers to extend their quantitative inquiry to include qualitative data and analyses to gain a more comprehensive understanding of the joint contributions of the multiple person-level factors (e.g., language skills, motivation) and contextual factors (e.g., peer relationships, perceived parental and cultural influences) theorized to contribute to reading comprehension development among culturally and linguistically diverse youth. As a result, the researcher can gain rich access to a broader range of factors contributing to the intervention outcomes observed - especially factors such as students' family context that, among marginalized populations, may be difficult to accurately and sensitively measure directly via quantitative approaches (Creswell & Plano Clark, 2011).

The current research study addressed these issues through a multi-phase, quasiexperimental mixed method evaluation of a reading intervention that targeted the development of adolescent vocabulary and reading comprehension skills of language minority and socioeconomically disadvantaged high school students. The overarching study goals were to: a) evaluate the efficacy of the Vocabulary Learning Project (VLP), an empirically-based, culturallysensitive reading intervention program, and b) explore person-level and contextual factors associated with differential response to the intervention. In order to address these goals, a multiphase mixed method design (Creswell & Plano Clark, 2011) was employed to allow for the combination of a quantitative study (Study 1) and a connected qualitative study (Study 2) with a third integration study (Study 3).

Specifically, the first phase of the research (Study 1) aimed to establish the effectiveness of the intervention (relative to a business as usual control condition) in improving students' reading comprehension and vocabulary, and to examine pre-intervention person-level factors that may differentiate outcomes (i.e., word reading skills, language comprehension skills, reading strategy use and motivation for reading) using quantitative methodology. Differential intervention effects on the basis of students' baseline language comprehension skills, reading strategy use and motivation for reading were assessed. Additionally, the relative roles of word reading skills, language comprehension skills and motivation for reading in the prediction of reading comprehension gains post-intervention were explored. In the second phase of the research (Study 2), various contextual factors potentially associated with differential responses to the intervention (e.g., peer, family, community and cultural influences on intervention experiences) were explored qualitatively. The final phase of the research (Study 3) involved an integration of findings from phases one and two, summarizing the overall potential of this intervention as a tool for reading skill improvement.

1.3 Overview of Studies

The primary objective of this study was to evaluate the outcomes of an empirically-based, afterschool reading intervention called the Vocabulary Learning Project (VLP), that targeted the improvement of the academic vocabulary and reading comprehension skills of academically atrisk high school youth in grades 9 to 11 within a disadvantaged community. The evaluation study (Study 1) employed a pretest-post-test control group design. The secondary goal of Study 1 was to explore in more depth the factors associated with students' response to intervention¹, including individual person-level (cognitive, linguistic, psychological) factors that potentially differentiated those who responded positively to the intervention from those who did not.

Broader contextual factors are typically not fully captured using purely quantitative data collection approaches (Creswell & Plano Clark, 2011). In fact, qualitative data collection approaches can be instrumental in elucidating broader contextual issues such as community or cultural influences that may be related to the outcome of interventions (Weine et al., 2005; Creswell & Plano Clark, 2011). Therefore, a qualitative strand of inquiry (Study 2) was sequentially embedded within the quantitative strand to allow for elaboration and enhanced interpretation of the findings from the quantitative impact study. In this way, the various data sources were able to inform each other. Specifically, the use of an embedded quasi-experimental mixed methods design allowed for the evaluation of intervention effects in addition to factors that potentially differentiated intervention responders from non-responders. The final sequential research phase (Study 3) involved an integration of findings across the prior two studies. Overall, this sequential multi-phase mixed method approach was adopted to inform an understanding of the program's potential as an agent of reading skill development and highlighted ways in which future intervention efforts may be optimally adjusted to maximize participant responsiveness.

1.4 Research Questions

The following research questions and associated hypotheses were addressed in this study.

Study 1

¹ It is important to distinguish that "response to intervention" in the context of this study specifically refers to participants' ability to demonstrate improved reading outcomes at immediate post-test in response to receiving the VLP intervention.

(1) What is the impact of an intensive after-school reading intervention on the vocabulary knowledge and reading comprehension skills of academically at-risk youth, when compared to a control (no intervention) condition?

It was hypothesized that youth in the intervention condition would demonstrate greater improvements in vocabulary knowledge and reading comprehension following exposure to the intervention, relative to participants in a control (business as usual) condition.

(2) What is the impact of the intervention on the following baseline subgroups of students relative to their control group counterparts : 1) students who are struggling (i.e., below average) with their language comprehension skills, and students who are not struggling (i.e., at least average), 2) students who report low levels of reading strategy use and students who report at least average reading strategy use, and 3) students who report at least average motivation to read and students who report low motivation to read?

It was hypothesized that students who participated in the 13-week VLP reading intervention and who started the intervention with a) relatively weaker language comprehension skills, b) low self-reported levels of reading strategy use, or c) greater self-reported motivation to read, would demonstrate significantly greater gains in vocabulary and reading comprehension relative to their business as usual control group counterparts (who also have weaker language comprehension skills, low reading strategy use and greater motivation to read).

(3) Among youth in the intervention group, are word reading skills, language comprehension skills, and motivation for reading (as assessed at pre-test), predictive of gains in reading comprehension skills above and beyond age, non-verbal ability and length of residence in Canada? It has been well established by prior research that factors such as age and non-verbal ability (used as a general measure of overall cognitive functioning) significantly account for a large proportion of variation in reading achievement (see McGrew & Wendling, 2010 for a comprehensive review). Among second language learners, length of residence in Canada (which often determines their age of initial exposure to the second language) has also been established to be an important factor accounting for their second language (L2) reading comprehension (e.g., Pasquarella, Gottardo & Grant, 2012). In this study it was important to acknowledge the role of these factors while assessing the additional roles of other person-level factors (i.e., word reading, language comprehension, and motivation to read) in predicting reading comprehension gains. Thus, it was hypothesized that after accounting for age, nonverbal ability and length of residence in Canada, poorer word reading and language comprehension skills in addition to higher motivation to read would be associated with greater gains in reading comprehension skills following the intervention.

Study 2

(4) Using the outcome data from Question 1 to identify and compare program responders (i.e., those whose reading outcomes improved) with non-responders (i.e., those whose reading outcomes did not improve) at the individual case level, what contextual factors are uniquely associated with responsiveness and/or non- responsiveness to the intervention from the perspective of participants?

Study 3

(5) How do findings from the qualitative study enhance findings from the quantitative study to facilitate understanding of the differential impact of the VLP intervention?

1.5 Mixed Methods Approach

This study involved a sequential embedded quasi-experimental mixed methods design consisting of two distinct yet interrelated phases (Creswell & Plano-Clark 2007). In the first phase, three sequential cohorts of youth from the Pathways to Education (P2E) program in Toronto's Regent Park community (see Chapter 2 for a complete description) were recruited over three consecutive years to participate in the intervention study. Each sequential cohort was conceptualized within a pretest-intervention-post-test framework. All participants were tested at pre- and post-intervention on a battery of outcome measures (i.e., reading comprehension and vocabulary knowledge) and theoretically-based predictors of outcomes (e.g., motivation for reading). Program fidelity was also monitored in this phase by the concurrent collection of observational data of every intervention session through audio-recordings of entire sessions. The aim of this phase of the study was to compare the intervention and control group participants' gains with respect to the intervention outcomes in order to ascertain the impact of the VLP (i.e., to determine whether it was effective in improving reading comprehension and vocabulary knowledge). This phase also informed program implementation (e.g., process, fidelity, student engagement) through observational data gathered from audio-recordings. Another goal of Study 1 was to determine whether certain person-level demographic, linguistic and psychological factors assessed at pretest would predict the outcomes and differentiate responders from nonresponders. This partially addressed the study's secondary objective of determining *person-level* factors associated with program outcomes that were measured quantitatively.

The second phase of the study (Study 2) also addressed the study's secondary objective by allowing for inductive exploration of the person-level *and* broader contextual factors that may differentially influence program impact. In this phase, participants were purposefully selected to provide in-depth interview accounts of their experiences in the intervention. These participants were then categorized as intervention *Improvers* or *Non-Improvers* based on their individualized program outcomes, as determined by results of the first phase which indicated whether or not each students' outcome scores improved. The goal was thus to determine whether there were differences in the thematic patterns emerging from interviews with *Improvers* compared to interviews with *Non-Improvers*, and ultimately identify factors differentially associated with youths' response to intervention.

This embedded quasi-experimental design allowed for the inclusion of a sequentially embedded qualitative strand of inquiry that addressed a secondary research objective within a broader program evaluation objective to allow for elaboration and enhanced interpretation of the findings from the primary impact study. In particular the qualitative phase allowed for examination of the multi-level contexts (Bronfenbrenner, 1979; 2005) potentially associated with differential response to intervention through exploration of the experiences of participants with certain kinds of outcomes (Creswell & Plano-Clark, 2007; 2011). In this design (Figure 1), quantitative and qualitative data interfaced at the design level, with the results of the qualitative (QUAL) strand being dependent on the preceding quantitative (QUAN) strand. The QUAL and QUAN strands were equally weighted in order to assess both program impact and context to maximally account for the potential of the program in improving the reading comprehension and vocabulary outcomes of academically at-risk immigrant and language-minority youth. This design is especially appropriate for evaluating a community-based reading intervention in a multicultural context because it can capture both the direct impact of the program as well as the interpersonal and contextual factors (e.g., home literacy experiences, neighborhood violence) that influence the intervention's effectiveness. The quantitative and qualitative data sources were integrated in a third phase in order to address the final research question which sought to achieve complementarity across all of the study's research questions. Figure 1 outlines the data

collection, analysis, and integration process following a sequential embedded quasi-experimental mixed method research design.



Figure 1. *MMR Model: Sequential Embedded Quasi-Experimental Design Note.* QUAN = quantitative; QUAL = qualitative

1.6 Overview of Chapters

This dissertation consists of three studies that collectively explored the general and differential impact of a 13-week manualized reading intervention on the reading achievement of disadvantaged and language minority adolescents. The first study consisted of a quantitative examination of the effectiveness of the intervention as well as predictors of outcomes. This first study is presented in Chapter 2. The second study was a qualitative exploration of the person-level and contextual factors that differentiated youth whose reading comprehension skills improved post-intervention from those whose skills did not improve. The second study is presented in Chapter 3. An overall mixed method integration of both studies is provided in Chapter 4. In the first chapter, the reasons for undertaking this research, overall mixed method approach and main research questions were described.

Chapter 2: Reading Comprehension Development in Academically At-Risk Adolescents: The Role of Intervention and Individual Differences (Study 1)

2.1 Introduction

2.1.1 Demographic Context

The Organization for Economic Co-operation and Development (OECD, 2009) stated that poverty and income inequality in Canada increased dramatically in recent years. A more recent report indicates that since the year 2000 very little progress has been made in reducing the poverty gap among the working-age population (OECD, 2018a). Particularly notable is the finding that the distribution of poverty is not equal across all social groups in Canada: "Aboriginal peoples, recent immigrants and racialized communities, and persons with disabilities all continue to face higher levels of poverty than others and are at higher risk of long-term poverty" (Citizens for Public Justice, 2012, p.4; Statistics Canada, 2013).

These demographic disparities are significant given that on average Canada receives on an annual basis approximately 250,000 newcomers coming from a variety of ethnic, cultural and home language backgrounds. Immigrants currently constitute over 20% of the overall Canadian population. The overwhelming majority of these new immigrants settle in large cities (Statistics Canada, 2016a). As a result, urban schools must deal with a very diverse population of students, including growing numbers of visible minority and language minority students who come from homes where the primary language spoken is not the language of schooling (e.g., English or French). The economic prospects of their parents are often bleak, given recent data indicating that Canadian immigrants earn significantly less than native-born individuals even after accounting for educational attainment, age and place of residence (OECD, 2018a).

Further accentuating these demographic inequalities are studies showing that disturbingly high numbers of youth new to Canada historically did not complete high school— "approximately 46% to 74% in some jurisdictions, while the rate for the general population is much lower at 12% to 25%" (Dei, 2008, p. 347). These statistics do not only apply to recent immigrants; studies also show that other immigrants, visible minority children and impoverished youth from inner-city communities are the most likely to drop out of school (Tilleczek, 2008, 2013). More recent data offer much promise with respect to high school graduation of immigrant youth in Canada. The 2011 National Household Survey (NHS) conducted by Statistics Canada indicates that across Canada as a whole, immigrants had higher high school and university completion rates than native born Canadians (Statistics Canada, 2013). Despite this, immigrant students in Canada continue to have lower reading scores than native Canadian students (Hou & Zhang, 2015).

The factors contributing to the risk of school dropout of this demographic group are complex and can be described from an ecological theoretical perspective that recognizes the influence of risk factors upon multiple systems progressing from those that are more proximal to more distal from the individual (see Bronfenbrenner, 1977 for a description of ecological systems theory; see Study 2 for a more detailed description). A diverse and extensive body of literature has addressed a variety of these risk factors (e.g., Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Rumberger, 2011; Archambault et al., 2017). Dei (2008), for example, suggests that many immigrant and racialized youth experience feelings of alienation from others in the process of attempting to fit in with a socially devalued identity. Dei (2008) also suggests that many youth in the Canadian context face problems of racialized and gendered poverty (especially among Somalis and Afghans), in addition to posttraumatic stress (specifically among students coming from war zones). They also often face many forms of discrimination. Another factor that has been shown to be related to academic risk is language-minority status. Languageminority status is defined broadly in this study as having a primary language other than English spoken at home, where such students may range from having limited to full proficiency in the

second language (i.e., English in the present context) (August & Shanahan, 2006). Being considered language-minority has also been attributed to reading underachievement among youth in Canada (Statistics Canada, 2013; Hou & Zhang, 2015). These factors all hold implications for the educational achievement of immigrant and racialized youth (Dei, 2008). Other researchers suggest that these broader societal issues impact school outcomes indirectly through their influence on student motivation and engagement with learning (August & Shanahan, 2006; Cumming & Geva, 2012a; Fredricks, Blumenfeld & Paris, 2004; Lesaux, Harris & Sloane, 2012; Malcolmson, 2001; Wilson, 2012).

A significant dilemma compounding the academic risk of recent immigrant language minority youth in particular is that they face the challenging task of simultaneously developing their English language proficiency while learning academic content (August & Shanahan, 2006). As will be discussed later, these youth often require six to seven years to acquire age and grade appropriate academic English literacy skills (Farnia & Geva, 2011; Cummins, 1991). In fact, concerns over the language and literacy skills and academic achievement of immigrant children are constantly expressed in countries like Canada that receive large numbers of immigrants (HRSDC, 2008).

The literacy underachievement among at-risk youth is well documented in the literature (Tilleczek, 2013). Underdeveloped literacy skills can lead to academic failure in students, which can in turn lead to dropout (August & Shanahan, 2006). The economic implications of leaving school prematurely are serious and impairing, as these school leavers are more likely to be unemployed (Cumming & Geva, 2012a; Snow & Biancarosa, 2003). Educators need ways to ensure that schooling meets the needs of the complex, diversified population of students they serve. An exploration of a broader range of instructional and learning models that are culturally responsive is needed (Dei, 2008). However, before developing instructional modifications that,

as in the case of the present study, aim to improve the reading achievement of culturally diverse immigrant and linguistic minority youth, the theoretical bases for reading development in this population must be delineated.

2.1.2 Theoretical Context

Empirical Bases of Reading Development

Reading skill development is influenced by multiple factors, including intra-individual, contextual and instructional factors. Reading skills that develop early on, well before commencement of formal education, include oral language skills, understanding of concepts of print, familiarity with print, knowledge acquisition and understanding of text structures (August & Shanahan, 2006). Gradually, and with the aid of formal instruction, skills in phonological awareness, letter knowledge, spelling and phoneme to grapheme correspondence develop. As these skills are acquired, children learn to integrate these skills with the type of learning necessary for obtain meaning from what is read, through the development of skills in vocabulary and reading comprehension strategies (August & Shanahan, 2008; National Reading Panel, 2000).

The National Reading Panel (2000) identified five components of reading development among monolingual readers, namely phonemic awareness, phonics, fluency, comprehension and vocabulary. The National Reading Panel's (2000) report represents to date the most comprehensive summary of research on reading instruction. Subsequently, to address the reading development of English language-learning (ELL)² children and youth, the National Literacy

² The present study defines ELL youth as those who have home languages other than English, and are often in the literature distinguished from language minority youth based on their much more limited development of English proficiency (August & Shanahan, 2008). Given the heterogeneity among these students in terms of English proficiency, in this study the terms ELL and language minority are used interchangeably; however the original term used in the referenced studies within this literature review will be kept, to maintain consistency with the original authors.

Panel (NLP) released a report in 2006 examining literacy learning among language minority children and youth. (August & Shanahan, 2006). The NLP (2006) report, based on a systematic review of the literature, examined the effectiveness of these five components of reading instruction as it pertains to ELLs. Findings indicated that the five basic aspects of reading development were equally important for ELL instruction; however effect sizes of interventions accounting for these factors were lower in these populations (Shanahan & Beck, 2006; Lesaux, Geva, Koda, Siegel & Shanahan, 2008). In order to have beneficial effects for ELLs, these interventions often required modifications (Snyder, Witmer & Schmitt, 2017). Additionally, it has been established that these important components of reading development are cumulative and developmental in nature, involving the gradual acquisition of increasingly complex skills (Geva & Wiener, 2015).

Theoretical Foundations for Reading Comprehension Development

Reading comprehension is a key developmentally relevant area of reading skill achievement that must be addressed at the secondary level. Reading comprehension development is a multicomponent process that requires several distinct but related language, reading and cognitive competencies (Lesaux et al., 2014). Various theoretical models have been applied to explain individual differences in reading comprehension ability. Gough and Tunmer's (1986) *Simple View or Reading (SVR)* model suggests that reading comprehension is a product of two skills: decoding and language comprehension. Decoding is a word-level reading skill involving word-recognition, phonemic awareness and pseudoword reading (Gottardo & Mueller, 2009; Gough & Tunmer, 1986). Language comprehension is a combination of vocabulary knowledge, syntactic and morphological awareness (Pasquarella et al., 2012; Geva, 2006).

Despite the strong evidence base indicating that decoding is an important predictor of reading comprehension (e.g., Farnia & Geva, 2013; Gottardo & Mueller, 2009; Nakamoto,

Lindsay & Manis, 2007), over time (i.e., beyond the primary grades) developmental changes occur among monolingual learners with respect to the relationships between decoding and reading comprehension. Specifically, as students get older, word-level reading skills such as decoding account for less of the variation in reading comprehension as compared to language comprehension (Catts, Hogan & Adolf. 2005). This finding is supported by studies involving second language learners who were exposed to the second language and received instruction in that language at an early age (e.g., in primary and middle school) (Lesaux & Kieffer, 2010; Geva & Farnia, 2012; Verhoeven & van Leeuwe, 2012). This is in large part due to the fact that at this developmental stage, word reading skills are more automatized and less predictive of individual differences in reading comprehension ability (Storch & Whitehurst, 2002). However, it is important to note that when language learners initiate exposure to the second language at later ages (e.g., in adolescence or secondary school), word reading skills remain important predictors of reading comprehension, as these skills have not yet become automatized (Pasquarella, Gottardo & Grant, 2012). In adolescence, the linguistic demands of texts read in academic settings increase, requiring adolescents to rely more on academic vocabulary knowledge for effective comprehension (Perfetti, 2007). Consequently, researchers working with adolescent populations have in part shifted their focus to the language comprehension component of the SVR.

The language comprehension aspect of the SVR involves components such as vocabulary, grammar, morphological and syntactic skills that contribute to reading comprehension (Cain, 2007; Kieffer & Lesaux, 2008; Low & Siegel, 2005). However, the vocabulary knowledge component of language comprehension has been consistently shown to be a stronger predictor of reading comprehension for adolescent language learners and nativeEnglish speakers (e.g., August, Carlo, Dressler & Snow, 2005; Farnia & Geva, 2013; Geva & Farnia, 2012; Ouellette & Beers, 2010; Proctor, Carlo, August, & Snow, 2005).

A well-established body of research demonstrates the strong relationship between vocabulary knowledge and reading comprehension among L2 learners (August, Carlo, Dressler & Snow, 2005; Biemiller, 1999; Droop & Verhoeven, 2003; Farnia & Geva, 2011; Geva, 2006; Hutchinson, Whiteley, Smith, & Connors, 2003; Lesaux, Lipka & Siegel, 2006; Proctor et al., 2005; Verhoeven, 2000, Verhoeven & van Leeuwe, 2008; Mancilla-Martinez & Lesaux, 2010; Verhoeven & van Leeuwe, 2011). The *lexical quality hypothesis* (Perfetti, 2007) is a theoretical framework that has been used to describe the relationship between word based skills and reading comprehension. It posits that rich lexical representations, which involve strong connections at the phonological, orthographic, semantic and syntactic levels, support the development of reading skills. Specifically, high quality lexical networks facilitate automatic word identification, which allows cognitive resources to be dedicated to comprehension (O'Connor, Geva & Koh, 2018; Pasquarella et al, 2012). Overall, this body of research demonstrates that vocabulary knowledge is a key aspect of language comprehension essential for reading comprehension.

Vocabulary development among at-risk youth

English language learners and their peers living in low income communities exhibit significant and persistent delays in English vocabulary knowledge (Buly & Valencia, 2002; Bialystok & Feng, 2009; Farnia, & Geva, 2011; Lesaux & Kieffer, 2010; Nakamoto et al., 2007). Vocabulary knowledge can be considered in terms of vocabulary breadth, which concerns the size of the lexicon, or vocabulary depth, which concerns the quality of semantic representations of the lexical entries (Ouellette, 2006; Townsend & Collins, 2009). Farnia and Geva (2011) tracked the vocabulary growth over time of in ELLs and monolingual students from grades 1 to 6. Findings indicated that, in spite of a consistent improvement over the 6 years, a persistent gap between the vocabulary knowledge of the ELLs and monolinguals was maintained.

ELLs are also at heightened risk for struggling with academic vocabulary specifically. English academic vocabulary is a form of vocabulary used in academic settings and in academic texts, particularly in the secondary grades, and is essential for academic success (Biemiller & Slonim, 2001; Corson, 1997; Cunningham & Moore, 1993; Nation & Kyongho, 1995; Scarcella, 2003; Townsend & Collins, 2009). This form of vocabulary is often complex and abstract, making it particularly challenging. Although ELL's basic cognitive-linguistic skills such as phonological and orthographic processing develop similarly to their monolingual peers, their English semantic knowledge is often less developed, resulting in lower vocabulary breadth and depth. For example, Jean and Geva (2009) compared the vocabulary depth and breadth over time of monolinguals and ELLs in grades 5 and 6 who attended school in Canada since grade 1. They found that the monolinguals continued to outperform the ELLs in terms of vocabulary breadth and depth. ELLs did not demonstrate grade level appropriate academic vocabulary, and were more likely to be familiar with academic words known by younger monolingual students. These findings suggest that ELLs would likely benefit from more systematic exposure to academic vocabulary as well as instruction that facilitates deep, semantic word knowledge (e.g., Townsend & Collins, 2009; Lesaux et al., 2014).

A Theoretical Model of Reading Comprehension Among Language Minority Youth

The difficulties language minority youth and their otherwise at-risk peers face in terms of vocabulary acquisition and reading comprehension suggests that a more nuanced model of reading comprehension is required to account for the complexities implicated. A more recent expanded SVR framework (Geva & Wiener, 2015; Yaghoub-Zadeh, Farnia & Geva, 2012) builds on the original Gough and Tunmer (1986) SVR model in several ways to incorporate

various additional intra-individual level factors as well as broader contextual factors that complicate reading comprehension skill development among second language learners. The model is presented in Figure 2. According to the model, reading comprehension is at the top of a developmental hierarchy (representing the ultimate goal of reading). The model summarizes a variety of person-level and broader contextual factors related to reading comprehension development.

Person-level factors. *Person-level factors* in reading comprehension development include all of the core aspects of the original SVR model (e.g., word- and text-based skills, language skills). The expanded SVR model also takes into account findings suggesting that reading comprehension depends not only on basic word reading skills, verbal working memory, and language skills, but also on the use of *higher-level comprehension strategies* such as inferencing, comprehension monitoring, and accessing text structure knowledge (Cain, Oakhill,& Bryant, 2004; Li & Kirby, 2014). The significant role of higher-level comprehension strategy use in reading comprehension development also suggests that explicit instruction in strategies related to these text-based components of reading comprehension could represent beneficial components of reading intervention. At the person-level, the expanded SVR model further incorporates reader characteristics beyond cognitive-linguistic processing skills, such as: command of native language vocabulary; psychological factors such as motivation and engagement; age; prior schooling experiences, and acculturation, for example. These person-level, reader characteristics interact with text-based characteristics to either enhance or hamper reading comprehension.

Broader contextual factors. Furthermore, the expanded SVR model acknowledges additional broader social, emotional and other contextual variables that have been theorized to impact reading comprehension. These include factors such as home literacy, school, instructional and community environments as well as relevant educational policies (Geva & Wiener, 2015).

In the context of the present study, this model provides a framework for examining the person-level and broader contextual factors that are important in understanding individual differences in reading comprehension and differential response to reading intervention among language minority and socio-economically disadvantaged learners. This model also provides a framework within which reading interventions such as the present one may be designed and evaluated.



Figure 2. An expanded framework for understanding reading comprehension in language-

minority learners

Note. Modified and reprinted from *Psychological Assessment of Culturally and Linguistically Diverse Children and Adolescents: A Practitioner's Guide* (p.34), by E. Geva & J. Wiener, 2015, New York, NY. Springer. Copyright (2015) by Springer Publishing Company. Reproduced with permission of Springer Publishing Company via Copyright Clearance Center.

Vocabulary instruction of at-risk youth

Of the many factors that can influence the trajectory of reading development, one key external factor is quality of instruction (August & Shanahan, 2008; Snyder et al., 2017). As previously noted, the available research suggests that vocabulary knowledge is an important component of reading comprehension skill development among language minority students (Carlo et al., 2004; Yesil-Dagli, 2011; Lervag & Aukrust, 2010; Jeon & Yamashita, 2014). As such, there is a need to develop and evaluate instructional techniques that can enhance the academic vocabulary and reading comprehension of these youth.

Although a large proportion of existing research examining vocabulary instruction has been conducted with monolingual learners at the elementary school level (National Reading Panel, 2000; Shanahan & Beck, 2006; August & Shanahan, 2008), there is considerable evidence that the same techniques identified for monolingual students benefit ELLs, including adolescents and middle-schoolers (e.g., Shanahan & Beck, 2006; Cumming & Geva, 2012; Townsend & Collins, 2009; Snow et al., 2009; Proctor et al., 2011). Based on the broader literature on general vocabulary instruction, there is considerable support for the utility of providing readers with the opportunity to learn new word meanings in a variety of contexts, and involving students in the practice and personalization of word meanings (Townsend & Collins, 2009). In addition, research has shown that multiple exposures to words, multiple opportunities to apply them in oral language and reading, pre-teaching of vocabulary words, semantic mapping, the use of technology, and group discussions have also been shown to be effective instructional practices (e.g., Biemiller & Boote, 2006; Snow et al., 2009; Hwang et al., 2015). Schools do not typically provide explicit vocabulary instruction or explicit reading strategy instruction to students (Biemiller & Slonim, 2001). To date, within the Canadian context very few studies have systematically examined the outcomes of intervention that incorporates these evidence-based

elements of instruction with at-risk language minority youth at the high school level (Cumming & Geva, 2012b).

Vocabulary and reading comprehension intervention with at-risk youth

Many of the studies that evaluate vocabulary interventions have involved monolingual students, and those that included culturally and linguistically diverse students involved middle-school-aged youth, in the fifth to eighth grades (i.e., August et al., 2009; Carlo et al., 2004; Dalton et al., 2011; Lesaux et al., 2010; Lesaux et al., 2014; Lubliner & Smetana, 2005; Snow et al., 2009; Townsend & Collins, 2009; Vaughn et al., 2009; Vaughn et al., 2011). These studies indicated significant treatment effects on curriculum-based outcome measures of word meanings taught. There were also significant effects reported on experimental measures of morphological awareness (e.g., Lesaux et al., 2010), metacognitive skills (Lubliner & Smetana, 2005), and reading comprehension (Carlo et al., 2004; Lesaux et al., 2010; Lubliner & Smetana, 2005). Lesaux et al. (2010) also reported significant effects on a standardized measure of reading comprehension.

There is a persistent dearth of substantial research showing large intervention effect sizes among middle school to high school students (Snyder et al., 2017). In 2017, Snyder, Witmer and Schmitt conducted a systematic review of intervention studies involving ELLs. Review of the 10 studies that met their established review criteria (e.g., experimental or quasi-experimental) indicated that all the studies involved instruction in one or more of the basic reading components previously identified, took place within school contexts, involved participants in grade levels ranging from K to Grade 8, and largely focused on Spanish-speaking ELLs within the United States of America (Cruz de Quiros, Lara-Alecio,Tong & Irby, 2010; Ebert, Kohnert, Pham, Rentmeester Disher & Payesteh, 2014; Ehri, Dreyer, Flugman & Gross, 2007; Mancilla-Martinez, 2010; Santoro, Jitendra, Starosta & Sacks 2006; Spycher, 2009; Tong, Irby, Lara-
Alecio & Koch, 2014; Townsend & Collins, 2009; Vadasy & Sanders, 2010; Vaughn et. al., 2006). Four of the ten studies from the Snyder et al. (2017) review reported large effect sizes for reading comprehension, but only among students in the lower grades, namely K to 2nd grade. With respect to vocabulary outcomes, four of the ten studies reported large effect sizes for vocabulary outcomes for students from K to 8th grade. Results of the Synder et al. (2017) review also indicated that interventions that include instruction in multiple components are more effective (e.g., includes both vocabulary and reading comprehension instruction). However the evidence is not definitive as to whether multi-faceted interventions are beneficial for older students when it comes to reading comprehension outcomes. In fact, multi-faceted interventions have been established to be less beneficial for improving vocabulary outcomes as such outcomes seem to benefit from vocabulary-focused intervention (Snyder et al., 2017).

In the Canadian context, Cumming (2012) highlighted a series of non-experimental studies evaluating the impact of a literacy tutoring intervention for a group of culturally and linguistically diverse high school students living in an urban, socio-economically underprivileged community (e.g., Won Jun & Watanabe, 2012; Lin, Ramirez, Wilson & Geva, 2012; Cumming & Geva, 2012b). Findings indicated that sustained and systematic instruction focused on literacy and language as well as explicit vocabulary strategy instruction represented beneficial interventional targets (Cumming & Geva, 2012a).

Overall, these intervention studies highlight the potential benefits of providing rich vocabulary instruction to culturally and linguistically diverse youth. However, based on the literature reviewed, there were only a few studies involving language minority students that employed an experimental design (e.g., August et al., 2009; Lesaux et al., 2014; Townsend & Collins, 2009; Vaughn et al., 2009). In fact over the last few decades, there have been no published experimental or quasi-experimental studies targeting reading development among high

school language minority students in the Canadian context. This is an important observation because in Canada, unlike the United States for example, there are two recognized official languages (i.e., English and French), with distinctive implications for studies of second language acquisition and education in a context of nationally institutionalized bilingualism (see Cummins & Persad, 2014 for a discussion of these issues). Additionally, differing immigration policies and differing patterns of academic achievement among immigrant Canadian youth relative to immigrant background students in the United States and most European countries (OECD, 2018b) highlights the potential utility of reading intervention studies of language minority youth in Canada specifically.

Further experimental or quasi-experimental research is needed that is large scale, includes adolescents in the secondary grades, employs multiple and varied measures of vocabulary depth and breadth, and considers the joint contributions of person-level psychological characteristics (e.g., motivation to read) and linguistic characteristics (e.g., language comprehension) to reading comprehension development among culturally and linguistically diverse youth. To date, there is a considerable dearth of intervention evaluations that acknowledge and consider these complex influences.

Response to Intervention as a Function of Individual Characteristics (Differential Response to Intervention)

The relative impact of instructional interventions varies as a function of the abilities and characteristics of the learner (Cronbach & Snow, 1977; Connor, 2011). Prior research in the area of literacy has demonstrated that the skills students bring to the task of reading influences the impact of a given instructional approach on their reading development (Connor, 2011; Lesaux et al., 2012). Research on such treatment-by-student interactions in the area of reading

comprehension is less well developed. In their evaluation study of a vocabulary intervention for middle school native speaking and language minority youth, Lesaux et al. (2014) found that their reading intervention had a greater impact on students whose home language was not English and for those who entered the intervention with underdeveloped vocabulary knowledge. Townsend and Collins (2009) evaluated the effects of vocabulary instruction in an after-school setting, and found that language minority students who had stronger receptive vocabulary skills showed greater response to the intervention, as reflected by their post-intervention performance on a measure of words taught. Another example of a study that assessed the differential impact of baseline vocabulary was O'Brien et al.'s (2014) evaluation of a family literacy intervention among children from grades K-3. The study results were suggestive of differential intervention effects based on differing pre-intervention vocabulary knowledge of the participants (i.e., those with lower vocabulary at pre-test had more post intervention gains). Other specific student characteristics or person-level predictors of reading comprehension development besides language minority status and pretest vocabulary knowledge have been implicated by prior research and require further investigation into their differential impact on reading comprehension as a result of intervention. These include characteristics such as motivation to read (Lesaux et al., 2012; Ivey & Broaddus, 2001; Wigfield, Eccles & Rodriguez, 1998) and metacognitive awareness of reading strategies (e.g., Follmer & Sperling, 2018; DeGennaro, 2018), particularly as assessed prior to receipt of intervention. Given the dearth of reading intervention research among adolescent language minority youth, it is important to examine dimensions of these variables on which there are empirically established reasons to believe that intervention effects may vary. Exploring intervention effects corresponding to different subgroups of language minority youth who bear certain potentially differentiating characteristics, or examining predictors of intervention outcomes, are some ways to evaluate differential effects (e.g., Lesaux

et al., 2014). In the present study, word reading, language comprehension (on the basis of vocabulary and morphology) and motivation to read were specific linguistic and person- level characteristics evaluated for their potential in differentiating intervention effects, as few studies have jointly examined these factors.

2.1.3 Summary and Rationale for Present Study

There is a persistent gap in the reading achievement of socially and economically disadvantaged youth who are native English speakers with immigrant parents, or are recent immigrant English language learners (August & Shanahan, 2006; HRSDC, 2008; OECD, 2018b). The reading development of these at-risk youth is complex and multi-determined. The expanded SVR framework (Geva & Wiener, 2015) provides a model that incorporates various person-level and contextual characteristics that are theorized to be predictive of reading comprehension, including vocabulary knowledge, metacognitive strategies for reading comprehension and motivation for reading amongst others. Recent research indicates that both language minority and native language speaking adolescents living in disadvantaged contexts require significantly more targeted support in the specific areas of academic vocabulary and reading comprehension. Intervention efforts have thus begun to espouse explicit instruction in these domains, especially for language minority and at-risk student populations (Nagy & Townsend, 2012). Although some available studies suggest that direct instruction targeted at vocabulary and reading comprehension skills is a worthwhile approach, the research base is not sufficiently developed to inform large scale efforts (Lesaux et al., 2014). Studies such as the proposed study are needed to fill this knowledge gap.

Despite a recent increase in intervention efforts targeting the literacy of culturally and linguistically diverse youth, very few experimental studies have been conducted with high

school-aged adolescents, and to this author's knowledge, no such evaluation studies have been done in the Canadian context. Given the complexity of reading comprehension development among culturally and linguistically diverse learners, this study evaluated an intervention designed to provide intensive instruction in the application of vocabulary learning and reading comprehension strategies to enhance youth's academic vocabulary and reading comprehension development.

Prior research indicates that intervention alone often represents only one of multiple potential factors affecting outcomes (Becker, Roberts & Voelmeck, 2003). That is, even in instances that overall intervention effects among participants may not be significant, interventions may still exert differential effects based on differing participant characteristics. Therefore, beyond investigating the intervention's overall impact, this study aimed to demonstrate the intervention's relative efficacy in order to inform the optimal match between instruction and students' characteristics. To explore differential intervention effects, subgroup analyses were conducted to determine whether there were subgroups of youth for whom the intervention was differentially effective. As well, the following baseline predictors of postintervention gains in reading comprehension were assessed after accounting for participants' age, length of residence in Canada and nonverbal ability: 1) word reading skills; 2) language comprehension skills, and 3) motivation to read. Given the dearth of experimental studies evaluating empirically-based reading interventions among culturally and linguistically diverse adolescents within the Canadian context, the present study represents a significant contribution to the existing evidence base.

2.1.4 Research Questions

The following research questions were addressed in the present study:

- What is the impact of an intensive after-school reading intervention on the vocabulary knowledge and reading comprehension skills of academically at-risk youth, when compared to a control (no intervention) condition?
- 2. What is the impact of the intervention on the following baseline subgroups of students relative to their control group counterparts:1) students who are struggling (i.e., below average) with their language comprehension skills, and students who are not struggling (i.e., at least average), 2) students who report low levels of reading strategy use and students who report at least average reading strategy use, and 3) students who report at least average motivation to read and students who report low motivation to read?
- 3. Among youth in the intervention group, are word reading skills, language comprehension skills, and motivation for reading (as assessed at pre-test), predictive of gains in reading comprehension skills above and beyond age, non-verbal ability and length of residence in Canada?

2.2 Methods

The following methods section details the community context within which this study occurred, describes the VLP intervention and its components, as well as the study design, participants, procedures, measures used and analyses conducted.

2.2.1 Community Context

Pathways to Education Program

The high school students recruited for participation in the reading intervention evaluated in this study were already enrolled in a broader, community-wide after-school tutoring and mentoring program administered by Pathways to Education Canada, at its Regent Park neighborhood site. Pathways to Education is a community-based charitable initiative that is supplementary to regular schooling, with the goal of buffering the impact of poverty and other systemic barriers to academic achievement that exist in society-barriers which serve to marginalize certain demographic populations of students, as described in the previous chapter (Cumming, 2012; www.pathwaystoeducation.ca). Specifically, the program aims to lower high school dropout rates and increase participation in postsecondary education among low-income youth. This is accomplished through the provision of various forms of support: 1) academic support, in the form of after-school tutoring available four evenings per week in all academic subjects, delivered by adult volunteers within multiple sites in the Regent Park community; 2) social support, through group mentoring (Grades 9-10), career or specialty mentoring (Grades 11-12), employment and internships, career guidance and post-secondary school application assistance; 3) *individualized* support offered by Student Parent Support Workers who serve as a liaison between home and school and offer personalized supports to students ranging from academic to socio-emotional needs, and 4) financial support in the form of scholarships, bursaries, bus tickets, lunch vouchers or other needed financial assistance to support students' completion of high school education (Fernando, 2011).

Regent Park Community

Youth living in Regent Park face many challenges with respect to remaining engaged in high school and pursuing post-secondary education. The demographic characteristics of the Regent Park neighborhood are commensurate with the need for a large-scale community initiative such as Pathways to Education.

Regent park is one of the most economically disadvantaged communities in the city of Toronto, and historically represents one of the largest public housing projects in Canada (Toronto Community Housing, 2007; Rowen, 2012), although current efforts are being made to revitalize it into a mixed-income community. The community faces numerous challenges relative to other communities in Toronto, Ontario, including: high unemployment rates (9.6% vs. 8.2% Toronto average); a significant proportion of low income families (42% vs 20% Toronto average), despite having comparable levels of educational attainment (e.g., 42% of individuals hold a Bachelor's degree or higher vs. 44% Toronto average), and contains a high proportion of single-parent families (29% vs. 21% Toronto average) (Statistics Canada, 2016b). In 2011 (at the time of study recruitment), 51% of Regent Park residents were immigrants and the majority (75%) of its residents were visible minorities (National Household Survey [NHS], 2011). Current data indicate that these proportions have remained relatively stable: 50% are immigrants, similar to the city average of 51%, and are disproportionately of a visible minority (i.e., 70% vs. 51% Toronto average). Primary visible minority or ethno-cultural groups were South Asian, Black, Chinese, Southeast Asian, and Filipino, respectively. Primary home languages were Bengali, Tamil, Mandarin, Vietnamese, Cantonese and Chinese, respectively (Statistics Canada, 2016b).

In light of the risk factors typically faced by demographic groups represented by the Regent Park community (as discussed in the previous chapter), it is reasonable then, that Pathways to Education aspires to help high school students in low income communities like Regent Park to complete high school and move on to post-secondary education by providing them with the various financial, academic, and social supports necessary for achieving positive outcomes.

2.2.2 The 'Vocabulary Learning Project' Intervention

As a result of previous research demonstrating the importance of reading interventions that target the improvement of students' academic vocabulary and reading comprehension skills (e.g., Farnia & Geva, 2011; Biemiller, 2009; Kamil, Mosenthal, Pearson, & Barr, 2000; Lesaux & Kieffer, 2010; Snow et al., 2009; Coxhead, 2000), the VLP reading intervention was conceptualized by a team of reading researchers with an expertise in second-language reading (Geva, Farnia, Chen, & Gottardo, 2011). Under their leadership a team of graduate students with background in reading research developed the empirically-informed, manualized intervention program that targeted weaknesses in academic vocabulary and reading comprehension strategies (Pasquarella et al., 2013).

The program provided direct and explicit instruction in academic vocabulary and reading comprehension strategies with the goal of closing the gap between the highest and lowest adolescent achievers. The program was implemented as a series of 13 weekly one- and a half hour tutor-led sessions, held after school in a small group setting (e.g., four to six students per group). Each of the 13 sessions was structured identically, beginning with an icebreaker, two reading passages used for strategy instruction, and two games to facilitate consolidation of learning.

Two complementary approaches to academic vocabulary learning were taught systematically: (a) semantic relationships taught through activities such as targeting words with multiple meanings and concept maps; and (b) roots and derivations taught through word families and focus on prefixes and suffixes (Pasquarella et al., 2013). Targeted reading comprehension strategies included identifying audience, type and purpose of texts, re-reading, paraphrasing, making predictions and inferences, planning, comprehension monitoring and activating prior knowledge. A summary of vocabulary and reading comprehension strategies is provided in Table 1. A detailed description of specific program strategies and related activities is provided in the program manual (see Pasquarella et al., 2013).

All program content was delivered by trained tutors -- adult volunteers who received intensive training (conducted by four of the five doctoral students who co-developed the program) over the course of two to three weeks, in preparation for the delivery of effective,

research-based instruction of vocabulary and reading comprehension strategies. In addition to providing these tutors with techniques to motivate and engage students, they were also trained to cultivate and maintain cultural competence in their work with the students. Given that tutor trainers were doctoral students in developmental and clinical child psychology, they were also able to provide tutors with training in strategies to effectively manage adolescent socioemotional and behavior issues when required. To facilitate increased attendance and engagement, all participants were provided the opportunity to accumulate points (for attendance, punctuality and participation) during each session, toward winning a grand prize via a draw held at the end of the program. In the interim, students were also able to receive a token prize whenever they accumulated a sufficient number of points at a pre-determined time-point during the course of the program.

Given the anticipated difficulties in recruiting and maintaining adolescent participants in an after-school academic intervention within a socio-economically underprivileged context, it was important to deliberately incorporate additional program elements to attract student participation and directly facilitate ongoing engagement. These included the provision of extrinsic motivators such as snacks, volunteer hours in exchange for study participation (to help students partially fulfill mandated high school volunteer requirements for the Ministry of Education of the province of Ontario), prizes, and financial remuneration for their time. Despite some equivocation about the ethics involved in offering these extrinsic motivators, prior research examining participants' reflections about their program experiences as they related to engagement indicated that these elements were valued by participants (Iwenofu, 2013).

Program design features that support the VLP's potential as an engaging, culturally responsive intervention. It has been previously established that methods of maximizing adolescent engagement are critical considerations when designing literacy programs (Guthrie &

Wigfield, 2000). The following are instructional elements of the VLP that were further aimed toward maintaining student engagement as well as cultural sensitivity, adaptability and responsiveness: (1) the intervention included collaborative learning activities that gave students an opportunity to meaningfully engage with one another, increasing group discussion and collaboration while impacting language development (Stahl & Nagy, 2006); (2) every session was centered around a short piece of text that included new vocabulary words for instruction. Participants were encouraged to contribute texts of interest to them on an ongoing basis (e.g., newspaper or web article). Texts provided or contributed by participants were of varying difficulty in terms of language content, and varied by genre (e.g., poetry, fiction, narrative, expository). Provision of such text-based learning opportunities was consistent with important, empirically validated elements of sound reading instruction (Gambrell, Malloy, & Mazzoni, 2011); (3) the intervention was developmentally sequenced, such that students would be engaged in activities that allowed for the incremental building of language skills over the course of the intervention (Lesaux et. al., 2012). This was accomplished through session by session, highly tutor-scaffolded skill building, culminating in student-led instruction of the strategies taught with guidance from the tutor (a technique known as reciprocal teaching) during the final two sessions of the program (Palincsar & Brown, 1984). The goal of reciprocal teaching was to increase students' autonomy while consolidating their emerging skills, a potentially important factor for supporting their motivation to remain engaged in the program (Meece, Anderman & Anderman, 2006). A summary of the 13 VLP program sessions and related activities is provided in Table 2.

Table 1

VLP Instructional Strategies and Activities

Vocabulary	Reading Comprehension
t	Identifying purpose, audience, and text types
Creation of word bank cards from daily readings	Activities: group discussion (semi-structured)
	Activating prior knowledge
	Activities: Know-Want-Learn (KWL) chart,
Word + definition (English and home language)	group discussion
	Making predictions
	Activities: group discussion
Synonyms	
	Comprehension monitoring
	Activities: group discussion; root word analysis; re-reading
Sample sentences	
	Paraphrasing
	Activities: notetaking, story map, graphic organizer
Word families	0
	Inferencing
Mental picture	Activities: using "wh-" and "how" questions
	Summarizing
	Activities: KWL chart, group discussion,
Mind maps	notetaking, story map, graphic organizer
Semantic maps	

Table 2

VLP Session Overview

Session	Instructional Goal	Direct Vocabulary Instruction Activities	Reading Comprehension Strategies	Closing Game
1	Building Rapport & Establishing Reading Interests	n/a	n/a	ice-breaker
2	Identifying Purpose/Audience/ Text Types & Making Predictions	vocabulary card	audience/purpose/text type, making predictions, choral reading, silent reading	stump the tutor
3	KWL for Before Reading	vocabulary review, lightening round (game), vocabulary card	audience/purpose/text type, activating prior knowledge, KWL, shared reading	20 questions
4	Question Generation	vocabulary review, password (game), vocabulary card	audience/purpose/text type, generating questions, silent	charades

			reading, strategy review	
5	Comprehension	vocabulary review, 20	strategy review,	10 words or
	Monitoring & "Click	questions (game), vocabulary	audience/purpose/text type,	less
	or Clunk"	card	"click" and "clunk',	
			comprehension monitoring using	
			checklist, independent reading	
6	KWL for During	vocabulary review, memory	audience/purpose/text type, KWL,	tree race
	Reading	(game), vocabulary card	shared reading, strategy review	
7	"Click or Clunk" (Par	vocabulary review, go fish	strategy review,	10 words or
	Deux)	(game), vocabulary card	audience/purpose/text type, "click" and "clunk',	less
			comprehension monitoring using	
			checklist, independent reading	
8	KWL for After	vocabulary review,	audience/purpose/text type, KWL,	charades
	Reading	lightening round (game), vocabulary card	shared reading, strategy review	
9	Paraphrasing	vocabulary review, 20	audience/purpose/text type,	tree race
		questions (game), vocabulary	paraphrasing, guided reading,	
		card	strategy review	
10	Paraphrasing AND	vocabulary review, password	audience/purpose/text type,	stump the
	Summarizing	(game), vocabulary card	paraphrasing, summarizing, team reading, strategy review	tutor
11	Inferencing &	vocabulary review, memory	audience/purpose/text type,	10 words or
	Questioning	(game), vocabulary card	inferencing, questioning, team reading	less
12 & 13	Reciprocal Teaching	vocabulary review,	reciprocal teaching, strategy	choice
		vocabulary game vocabulary	review, before/ during/after	
		card	reading strategies, team reading	

2.2.3 Research Design

The design was quasi-experimental, with recruitment occurring annually across four cohort sequential cohorts between September 2011-June 2014 (i.e., 4 cohorts, multiple measurements, over 4 years). Participants in each sequential cohort were assigned to a pre-test/ intervention/ post-test group or pre-test/business-as-usual/ post-test group. In general, participants who were informed and interested in the study were first placed in the intervention group, and after intervention participants were at capacity, additional students were recruited to a control group, with the intention that they would serve as waitlisted controls who would ultimately participate in a subsequent intervention cohort. However, across the four cohorts of recruitment and study participation, none of the waitlisted control group students participated in

a subsequent wave of intervention (despite invitation to participate). Thus, there was no waitlisted control group.

2.2.4 Participants

The youth recruited for participation in this study were high school students in Grades 9 to 11, who were enrolled in the Pathways to Education Program – Regent Park site. Participants were initially identified and sometimes recruited by Pathways to Education staff directly liaising with the intervention research team. Several graduate student project coordinators were also responsible for recruitment and overall study coordination. Written informed consent was obtained from the parents of all participants under the age of 16, and from the participants directly if they were over the age of minority. Verbal and written informed assent was obtained from minors under age 16. All study procedures were approved by the University of Toronto Research Ethics Board.

Recruitment

Conducting community-based educational research can be a complex endeavor, particularly among marginalized, socio-economically underprivileged populations. As anticipated, in spite of the extrinsic motivators for participation provided (e.g., volunteer hours, financial remuneration) as well as direct support from Pathways to Education program staff in the identification and recruitment of potential participants, there were significant challenges encountered in overall participant recruitment and retention. The VLP project team (doctoral students and primary investigators) articulated some plausible sources of these recruitment challenges: 1) at the individual student level, poor motivation, perceptions of the researchers as community "outsiders", negative attributions of the study objectives, and study refusal due to perceived or actual pre-existing learning problems; 2) at the contextual level, factors associated with community demographics such as the fear of occurrence of neighborhood violence while attending the VLP after school, stress, competing priorities (e.g., preference for after school employment rather than study participation) and culturally-imposed limitations (e.g., avoidance of the likely close interaction with members of the opposite sex if placed in the same small group for weekly program sessions) (Azimi, Iwenofu, Fraser & Geva, 2013). Additionally, these participants were already voluntarily enrolled in the broader Pathways to Education intervention, which required them to attend after-school academic tutoring as previously described, at least twice a week. As such, it was possible that many potential participants declined participation given that it would require commitment to an additional academic intervention. For these reasons, recruitment and assignment to either study condition occurred in a sequential manner that maximized student enrolment in the intervention group followed by the establishment of an adequate business-as-usual control group sample.

Cohort 1. After the first round of recruitment, 48 participants were enrolled in the intervention group and completed the intervention. There were no participants assigned to the control condition.

Cohort 2. During the subsequent second round of recruitment 62 participants were initially assigned to the intervention condition. Intervention completers are defined as participants who completed at least 9 of the total 13 sessions of the intervention. Of the 62 who initially were enrolled, 17 participants did not complete the intervention (i.e., they completed more than two sessions but less than 9 of 13 sessions). Of the 17 non-completers, 9 completed only two sessions or less, and were considered eligible to serve as control participants. Of the 45 remaining intervention completers, 4 participants did not return for post-testing, resulting in a total of 41 intervention completers in this cohort. In total, there were 27 participants initially

assigned to the control group for cohort 2. Of these, 4 did not return for post-testing, leaving a total of 32 control participants in the 2^{nd} cohort.

*Cohort 3.*During the third round of recruitment (cohort 3), 47 participants were initially assigned to the intervention condition. Of the 47 who were initially enrolled, 13 were non completers. Of these, 5 were moved to the control group (as they completed 0-2 sessions), and 8 dropped out after completing a range of 3-8 sessions. Of the 34 remaining intervention completers, one did not complete post testing, resulting in a total of 33 intervention completers in this cohort. In total, there were 22 control participants in cohort 3, including waitlisted control and participants who met exclusionary criteria for the intervention group. Ten of these students did not return for post testing, resulting in a total of 12 control participants in this cohort.

Cohort 4. Given the poor recruitment of control participants in the study overall, a final cohort of 47 control participants was recruited, of which 9 did not return for post-testing -- resulting in a total of 38 control participants for this cohort. A series of flow charts depicting participant recruitment efforts across cohorts is provided in Figure 3.



Figure 3. Flow Diagram of Participant Recruitment









2.2.5 Measures

Questionnaires

Demographics. The Family Background Questionnaire was developed for use in this study (Appendix B). The questionnaire was completed by students in consultation with their parents, and provided demographic information on age, sex. native language, parental education, employment, country of origin, years in Canada, housing, who lives in the household, attendance in heritage programs, home languages and so on. This measure was administered at pre-test.

Motivation for Reading (MRQ). The Motivation for Reading Questionnaire (MRQ) is a 53-item, four-point Likert scale questionnaire (Wigfield & Guthrie, 1997) that was used to assess the extent to which each student was motivated to read (e.g., I like to read about new things). Responses on this measure range from 1 = very different from me to 4 = a lot like me, and scores across all items are averaged to yield a total score reflecting the student's extent of reading motivation. The total score has documented reliability (Cronbach's $\alpha = 0.43$ to 0.81) (Wigfield & Guthrie, 1997). Evidence for factorial validity of the MRQ has been established (Wigfield & Guthrie, 1997) and cross-validated (Unrau & Schlackman, 2006). This measure was administered at pre- and post-test.

Metacognitive Awareness of Reading Strategies. The Metacognitive Awareness of Reading Strategies Inventory (MARSI) is a 30-item, five-point Likert scale self-report questionnaire (Mokhtari & Reichard, 2002) that was used to assess the students' metacognitive awareness and perceived use of reading strategies when reading academic or school-related material (e.g., I decide what to read closely and what to ignore). Responses range from 1=I*never or almost never do this* to 5=I *always or almost always do this*. The measure comprises three subscales (i.e., Global Reading Strategies, Problem Solving Strategies, Support Reading Strategies) and a total score. An average score is computed for each subscale as well as a combined average score ranging from 1 to 5, indicating one of three levels of reading strategy use per student: high (mean of 3.5 or higher), medium (mean of 2.5 to 3.4), and low (2.4 or lower). The measure has good overall reliability (Cronbach's alpha = .89), as well as reliability coefficients ranging from 0.89 to 0.93 across subscales and grade levels from 6th to 12th grades. The measure also has good construct validity (Mokhtari& Reichard, 2002). This measure was administered at pre- and post-test.

Non-verbal cognitive ability

Nonverbal Ability Matrix Analogy Test (MAT – Expanded Form, Naglieri, 1985). The MAT was used to assess student's nonverbal reasoning skills. This measure was individually administered as a cognitive control measure that is relatively free of language and culture related confounds, and is typically used in research involving culturally and linguistically diverse second language learners. For each item in this measure, participants were presented with an incomplete visual pattern and asked to select one of six options to complete the pattern. The test consists of four subtests, each with 16 items for a total of 64 items from which a total score, standard score and percentile score was derived for each participant. The MAT has high internal consistency reliability (Cronbach's α = .88 to 0.95) and good criterion validity (Harris, 1988). This measure was administered at pre-test.

Vocabulary

Expressive vocabulary. Expressive vocabulary in English was assessed using the Expressive Vocabulary Test, Second Edition- Form A (EVT-2, Williams, 2007), a 190-item individually administered standardized test. For each item, the examiner presents a picture and reads a stimulus question to which the participant is required to respond with one word that provides an acceptable label, answers a specific question, or provides a synonym for a word that matches the picture appropriately. A raw score, total score and percentile score are calculated per

participant. The EVT-2 has good internal consistency reliability (Cronbach's α =0.94) as well as good construct and content validity (Williams, 2007). This measure was administered at pre- and post-test.

Receptive vocabulary. The Peabody Picture Vocabulary Test-Fourth Edition- Form A (PPVT-4) (Dunn & Dunn, 2007) measured receptive vocabulary in English. During this individually administered standardized test, the participant hears a word read-aloud, is shown four pictures and then asked to select the picture that best matched the orally presented word. There is a maximum of 228 items from which a total raw score, standard score and percentile score is derived per participant. The PPVT-4 has good internal consistency reliability (Cronbach's α =0.94) as well as good construct and content validity (Dunn & Dunn, 2007). This measure was administered at pre- and post-test.

Academic vocabulary. Definitions of academic root words (e.g., arm, garbled) mastered at different grade levels was examined with selected items from the Biemiller Root Word Inventory (Biemiller & Slonim, 2001). For each of the 30 items, a sentence was orally read to participants, after which the participants were required to provide the definition of a single word within that sentence (i.e., target vocabulary words were embedded within sentences). Responses could be indicated verbally, by pointing or by acting to explain the meaning of the word. The 30 items were presented in order of difficulty beginning with the least difficult (Level 2) words to the most difficult (Level 12) words. Scores could range from 0 to 30. A total raw score was derived for each participant. This measure has been shown to correlate highly with PPVT picture items (r=.79) and PPVT context sentence items (r=.86) (Biemiller, 2005). This measure was administered at pre- and post-test.

Words with multiple meanings (vocabulary depth). The Words with Multiple Meanings (WMM) task was used to assess academic vocabulary depth. This 21-item, multiple choice task

was developed for use in this study, and assesses familiarity with multiple word meanings. Items are comprised of selected words and contextual sentences from Biemiller (2009)'s word list. Biemiller's (2009) list was constructed such that each word on the list was consistent with empirical evidence about the kind of vocabulary essential for acquisition by students (including English language learners) in educational settings (Biemiller, 2012). For each item, students were asked to read a sentence with a bolded target vocabulary word and asked to select the meaning of the target word from three choices (e.g., "Cats are included with a **class** of animals called mammals: In this sentence class means: a) group of like plants or animals; b) group within a school; c) kind or sort)". A total raw score was calculated for each participant. In this sample, the WMM total score was moderately correlated to raw scores on other measures of vocabulary such as the PPVT-4 (r=0.69) and EVT-2 (r=.57). This measure was administered at pre- and post-test.

Morphological Structure

Test of morphological structure. The Carlisle Test of Morphological Structure (TMS-Carlisle, 2000) is a 28-item experimental task that assesses awareness of the relations of root and derived word forms. For each item, participants are asked to use a target stimulus word in a sentence by modifying the word to appropriately fit the sentence (e.g., *major*. He won the vote by a ... [*Majority*]). A total raw score was derived for each participant based on the number of items answered correctly. Validity evidence has been supported based on evidence of moderate to high correlations with other measures of morphology (r=0.40 to 0.64) (Carlisle, 2000). This measure was administered at pre- and post-test.

Word-Level Reading

Word-level reading. The Sight Word Efficiency (SWE) and Phonetic Decoding Efficiency (PDE) subtests of the Test of Word Reading Efficiency-2nd Edition (TOWRE-2;

Torgesen, Wagner, & Rashotte, 2012), an individually administered standardized test, were used to assess word reading accuracy. For both the SWE and PDE subtests, participants were asked to accurately read aloud a list of real words and non-words respectively, of increasing length and difficulty. The two subtests combine to form a Total Word Efficiency (TWRE) index. Test-retest reliabilities of the TOWRE-2 range from .89 to .93 across the subtests. Index scores on the measure have correlations ranging from 0.89 to 0.96 with similar measures of achievement (Tarar, Meisinger & Dickens, 2015).

Text-Level Reading: Reading comprehension

Gates MacGinitie Reading Test. The comprehension subtest of the Gates MacGinitie Reading Test – Forms S and T ([GMRT], MacGinitie, MacGinitie, Maria, & Dreyer, 2000), is a norm-referenced, timed 45-item multiple-choice reading comprehension test that can be administered in groups. Participants were asked to read short passages and afterward answer 3-5 comprehension questions. Reliability coefficients are in the range of 0.90 for forms S and T, and both forms have high inter-correlations ranging from 0.8 to 0.85 (MacGinitie et al., 2000). Form S was administered at pre-test, and the alternate form T was administered at post-test.

Woodcock Reading Mastery Test. The passage comprehension subtest of the Woodcock Reading Mastery Test-Revised ([WRMT-R], Woodcock, 1987) was used to assess reading comprehension. It is a modified cloze task that requires participants to fill in one or more blanks per item according to the surrounding phrases or sentences by selecting the appropriate missing word from a multiple choice array. A total raw score, standard score and percentile score was derived per participant. The WRMT-R has good internal consistency reliability (coefficients exceed $\alpha = 0.80$) and well established validity. It was selected as a primary outcome measure in this study given its wide use in remedial outcome research, and was administered at pre- and post-test.

2.2.6 Preliminary data analysis

Missing data

Over the four years of data collection, there was inevitable attrition in the overall sample, in part due to loss to follow-up of participants across cohort years. There was 10.28% missing data as a result of attrition. To handle the loss of sample size and avoid biases in estimation, a multiple imputation procedure (Rubin, 1987; Little & Rubin, 2002) was used to estimate incomplete data points. Multiple imputation is one of the optimal methods for managing missing data, and provides unbiased parameter estimates (Schafer & Graham, 2002). First, in order to verify that overall attrition did not systematically contribute to variability in the key outcome variables (i.e., vocabulary and reading comprehension), as well as to differences in nonverbal ability, gender, and age of participants, attrition bias was examined for the entire sample. Independent samples t-tests (with a Bonferronni correction) and a chi-square test (for gender) indicated that there were no statistically significant differences among study participants and dropouts with respect to these factors, p > .05 (see Table 3 for a summary of mean scores). Subsequently, the missing data was multiply imputed (m=5) through the linear regression method. This resulted in a final sample of 232 participants with complete data for analyses (127) participants were in the intervention group, and 105 participants were in the control group). The data imputation did not significantly alter the overall mean of the measures for which missing data was imputed. All subsequent analyses were conducted using the five resultant data sets, which were combined using appropriate procedures for aggregation of imputed data (Schafer & Graham, 2002; Rubin, 1987) such as applying the same statistical analysis to each of the imputed data sets and then combining the analyses into one pooled result so that uncertainties due to the missing data can be accounted for (van Ginkel & Kroonenberg, 2014). Notably, results from all

subsequent analyses were similar when conducted individually with the multiply imputed data

sets or when conducted with the reduced data set of students with complete data.

Table 3

Characteristics of Study Participants and Dropouts

	Participants (N=191)		Dropouts (N=57)	
Characteristic (Measure)	n (%)	M (SD)	n (%)	M (SD)
Gender				
Female	123 (64.4)		27 (47.37)	
Male	68 (35.6)		30 (52.63)	
Age		15.11 (1.13)		15.85 (1.36)
Nonverbal ability (MAT)		51.40 (10.82)		49.12 (11.31)
Baseline vocabulary				
Expressive vocabulary (EVT-2)		123.99		121.42 (16.92)
-		(21.61)		
Receptive vocabulary (PPVT-4)		167.89		165.26 (24.42)
		(27.52)		
Baseline reading comprehension				
Passage Comprehension		43.97 (8.35)		40.67 (6.22)
(WRMT-R)				
Reading Comprehension		25.14 (10.36)		21.35 (8.73)
(GMRT)				

Note. Raw scores were used. MAT = Matrix Analogies Test; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests; For all t-tests, p>0.05.

Attrition of Participants Within Intervention Condition

To account for attrition bias within the intervention group, the demographic

characteristics of intervention completers (n=127) and non-completers (n=16) were compared to

determine whether non-completers could be dropped from the analysis without distorting the

representativeness of the sample. Chi-square tests (and Fisher's Exact tests when indicated) were

used to compare discrete demographic variables, and independent samples t-tests were used to

compare continuous baseline characteristics. Independent samples t-tests (with a Bonferronni

correction) indicated similar representation among intervention completers and non-completers for age, age at arrival in Canada, as well as baseline (pre-test) performance on tasks assessing nonverbal ability, word reading, vocabulary knowledge and reading comprehension, p > .05. Chi-square and Fisher's Exact tests indicated similar representation for demographic variables, p> .05. These included gender, country of birth, home language and maternal education (Table 4). As a result of these analyses, it was concluded that the overall representativeness of the sample would be preserved after excluding the 16 intervention non-completers from the study analyses. Therefore, within the entire sample of intervention group participants, only the 127 intervention completers were included in the analyses.

Table 4

Characteristics of Intervention Completers and Non-Completers

	Completer	rs (N=127)	Non-Cor (N=16)	npleters	Chi square tests of independence/Fisher's
					Exact test/Independent
Characteristic	n (%)	M(SD)	n (%)	М	Samples t-test
(Measure)	II (70)	M(5D)	II (70)	(SD)	
Gender					$X^{2}(1)=.52,$
					<i>p</i> =.47
Female	83 (65.4)		8 (50.0)		
Male	44 (34.6)		8 (50.0)		
					<i>t</i> (141)=-
Age		15.32 (1.26)		14.82(0.92)) $1.25, p=.21$
Country of Birth					$X^{2}(1)=.10,$ p=.76
Canada	40 (55.6)		5 (50.0)		I
Outside Canada	32 (44.4)		5 (50.0)		
Age at arrival in Canada		8.24 (4.76)		9.00 (9.90)	
					Fischer's
Home language					exact <i>p</i> =.16
English	33 (75.0)		6 (75.0)		
Other	11 (25.0)		2 (25.0)		
					Fischer's

Maternal Education					exact <i>p</i> =.66
Secondary	31 (43.7)		1 (12.5)		
level or below					
Postsecondary	17 (23.9)		3 (37.5)		
level or above					
Not indicated	23 (32.4)		4 (50.0)		
Nonverbal		51.50 (11.10)		49.20 (11.13)	t(141)=-
ability (MAT)					1.15, <i>p</i> = .25
Baseline word					
reading					
TOWRE –		82.91 (12.96)		83.11 (5.09)	t(141)=.05,
word reading					<i>p</i> = .96
TOWRE –		43.60 (12.76)		43.10 (11.06)	t(141) =12,
pseudoword					<i>p</i> = .91
reading					
Baseline					
vocabulary					
Expressive		123.40		116.38 (22.55)	t(141) =72,
vocabulary		(23.31)			<i>p</i> =.48
(EVT-2)					
Receptive		164.26		168.22 (19.54)	t(141)=.73,
vocabulary		(31.63)			<i>p</i> =.47
(PPVT-4)					
Baseline reading					
comprehension					
Passage		43.69 (9.60)		39.30 (5.70)	t(141) = -
Comprehension					1.32, <i>p</i> = .19
(WRMT-R)					
Reading		23.92 (10.72)		21.90 (10.59)	t(141) =57,
Comprehension					<i>p</i> = .57
(GMRT)					

Note. MAT = Matrix Analogies Test; TOWRE = Test of Word Reading Efficiency; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests; For all t-tests, p>0.05.

Characteristics of recruitment cohorts at baseline

Given that participant recruitment and study procedures (e.g., intervention program

implementation) occurred sequentially over four cohort years, it was important to evaluate

whether there were any systematic differences between participant cohorts at baseline, as well as

to examine intervention implementation across the three intervention cohorts to ensure that cohorts could be combined into one aggregate sample for subsequent analyses.

Intervention group cohorts (cohorts 1-3). A chi-square test of independence was used to determine whether there were any significant gender differences between the three cohorts of intervention group participants [(i.e., cohorts 1 (n=48), 2 (n=45) and 3 (n=34)]. Results indicated that there were no significant gender differences ($X^2 = 2.042$, df= 2, p=.360). A one-way analysis of variance was conducted to determine whether the three intervention cohorts differed significantly with respect to age, age at arrival in Canada, nonverbal ability, and baseline performance on multiple measures of vocabulary and reading comprehension. Results indicated that there were no significant differences between the intervention group cohorts with regard to nonverbal ability [F(2,124)=2.24, p>.05] and age at arrival in Canada [F(1,29)=0.74, p>.05]. There was a statistically significant difference in age [F(2,124)=6.50, p=.002]. A follow up posthoc Tukey test was conducted to evaluate pairwise differences in mean age across the intervention cohorts. Intervention group participants in cohort 1 were significantly older at baseline (M=15.81, SD=1.25) in comparison to those in cohorts 2 (M=15.1, SD=1.32) and 3 (M=14.93, SD=.94). There was also a statistically significant difference between the cohorts on measures of expressive and receptive vocabulary [EVT-2: F(2,124)=9.96, p<0.001) and PPVT-4: (F(2,124)=7.45, p=.001 respectively]. Post-hoc Tukey tests indicated that participants in cohort 1 scored significantly lower on these measures [EVT-2: M=111.01, SD=21.36; PPVT-4: M=150.23, SD=31.59], compared to participants in cohorts 2 [EVT-2: M=131.0, SD=24.10; PPVT-4: M=171.74, SD=31.21] and 3 [EVT-2: M=127.32, SD=23.60; PPVT-4: M=172.31, SD=29.59] respectively. On other measures of vocabulary knowledge, cohort 1 intervention participants also scored significantly lower than cohort 2 and 3 at baseline (see Table 5). There was a statistically significant difference between the three intervention cohorts on two measures

of reading comprehension [WRMT-R: F(2,124)=7.54, p=.001; GMRT: F(2,124)=17.97, p<.001]. Post-hoc Tukey tests indicated that cohort 1 participants scored significantly lower on average, on these measures at baseline (see Table 5 for a summary of the distribution of scores across cohorts on these measures).

Table 5

Characteristics of Intervention Cohorts (Cohorts 1, 2 and 3)

	Cohort 1 (n=48)	Cohort 2 (n=45)	Cohort 3 (n=34)
Characteristic (Measure)	M (SD)	M (SD)	M (SD)
Age	15.81(1.25)***	15.1(1.32)	14.93(.94)
Age at arrival in Canada	9(3.98)	7.53(5.42)	8.24(4.76)
Nonverbal ability (MAT)	50.23(12.63)	49.9(11.85)	54.82(9.14)
Baseline vocabulary			
Expressive vocabulary (EVT-2)	111.01(21.36)***	131 (24.1)	127.32(23.6)
Receptive vocabulary (PPVT-4)	150.23(31.59)***	171.74(31.21)	172.31(29.59)
Academic vocabulary (BWMM)	15.12(3.71)**	16.58(3.88)	17.32(2.20)
Vocabulary (BRWI)	12.18(4.81)*	14.64(4.77)	15.38(4.71)
Baseline reading comprehension			
Passage Comprehension	39.5(8.74)***	44.12(8.53)	47.45(11.25)
(WRMT-R)			
Reading Comprehension	17.11(7.19)***	27.17(10.69)	28.29(11.33)
(GMRT)			

Note. Raw scores were used. MAT = Matrix Analogies Test; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; BWMM = Biemiller Words With Multiple Meanings; BRWI= Biemiller Root Word Inventory WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests. *p<.05; ** $p\leq01$; *** $p\leq.001$

Program implementation across intervention cohorts. In order to combine all

intervention cohorts for subsequent analyses, it was important to examine intervention program

implementation across cohorts to verify whether there were any significant variations in

implementation that may bias results when cohorts are aggregated. In order to evaluate program

implementation across cohorts, the two different versions of the program manual used across cohorts, as well as tutor training procedures and program adherence rates for each cohort were compared (see Appendix A for a detailed description and results of these analyses). Results indicated that there were significant differences in program implementation and sample composition within cohort 1 which could have potentially biased the results of this study, if cohort 1 data were retained in subsequent analyses. Specifically, results of those analyses highlighted the following differences in program implementation in cohort 1. During this first cohort, a pilot version of the intervention manual was used (version 1). During the second and third intervention cohorts, a slightly revised manual was implemented (version 2). For the purposes of this study, the intervention as implemented in the first cohort (version 1) was considered to represent a different, pilot version of the intervention. Version 1 of the intervention was also conducted with a sample of youth in cohort one who were, on average, identified and referred to participate in the intervention condition on the basis that they were struggling readers (based on a review of documented communication logs within the first year of program implementation). Sampling differences between the cohorts were further evidenced by statistically significant differences between cohort 1 and the other two cohorts on pretest measures of vocabulary and reading comprehension as previously described. Given that the goal of this study was to evaluate the reading intervention as operationalized in the most finalized version of the program manual (i.e, version 2, as employed with cohorts 2 and 3), with equal opportunities for participation from students of varying baseline vocabulary and reading comprehension skills, data from cohort 1 participants were removed from subsequent analyses. Data from participants in cohorts 2 and 3 of the intervention group were then aggregated to form the final sample of intervention participants (n=79). It should be noted that outcome analyses for treatment effects were also conducted with data from intervention cohort 1, and revealed similar

outcomes to analyses with the entire sample of intervention participants (i.e., cohorts 2 and 3) across all relevant outcome measures (see Appendix A for a summary).

Control group cohorts (cohorts 2-4). A chi-square test of independence was used to determine whether there were any significant gender differences between the three cohorts of control group participants [(i.e., cohorts 2 (n=36), 3 (n=23) and 4 (n=46)]. Results indicated that there were no significant gender differences ($X^2 = 8.47$, df= 2, p=.076). A one-way analysis of variance was conducted to determine whether the three control cohorts differed significantly with respect to age, age at arrival in Canada, nonverbal ability, and baseline performance on multiple measures of vocabulary and reading comprehension. There was a statistically significant difference in age [F(2,102)=6.50, p<.001)]. Follow-up tests indicated that control participants in cohort 2 were significantly older at baseline (M=16.12, SD=1.07) in comparison to those in cohorts 3 (M=15.51, SD=1.18) and 4 (M=14.41, SD=.60). Cohort 3 controls were significantly older than cohort 4 controls. Additionally, analysis of variance results indicated that there were no significant differences between the control group cohorts with regard to nonverbal ability [F(2,102)=1.47, p>.05], age at arrival in Canada [F(2,17)=0.58, p>.05], and baseline performance on measures of vocabulary knowledge and reading comprehension [e.g., EVT-2: F(2,102)=1.42, p>.05), PPVT-4: F(2,102)=0.42, p>.05); WRMT -R: F(2,102)=1.45, p>.05)], with the exception of one reading comprehension measure [(GMRT: F(2,102)=3.56, p=.035)]. A post-hoc Tukey test indicated that cohort 2 control participants scored significantly lower than cohort 3 and 4 controls on the GMRT assessment of reading comprehension skills (see Table 6 for a summary of the distribution of scores across control cohorts across the measures used). Overall, data from participants in all control cohorts were aggregated to form the final sample of control group participants (n=105).

Table 6

Characteristics of Control Cohorts (Cohorts 2, 3 and 4)

	Cohort 2 (n=36)	Cohort 3 (n=23)	Cohort 4 (n=46)
Characteristic (Measure)	M (SD)	M (SD)	M (SD)
Age	16.12(1.07)***	15.51(1.18)***	14.41(.60)
Age at arrival in Canada	4(2.45)	6.17(4.36)	4.5(3.85)
Nonverbal ability (MAT)	49.28(11.58)	53.83(10.92)	49.72(10.02)
Baseline vocabulary			
Expressive vocabulary (EVT-2)	123.71(20.29)	129.39 (13.67)	122.28(15.76)
Receptive vocabulary (PPVT-4)	169.85(20.99)	174.07(17.83)	169.84(15.17)
Academic vocabulary (BWMM)	16.56(2.18)	17.88(1.88)	16.97(2.75)
Vocabulary (BRWI)	13.75(3.48)	15.45(3.10)	13.98(3.30)
Baseline reading comprehension			
Passage Comprehension	41.83(5.80)	44.43(5.00)	43.25(6.49)
(WRMT-R)			
Reading Comprehension	21.77(8.99)*	27.33(8.41)	26.48(10.00)
(GMRT)			

Note. Raw scores were used. MAT = Matrix Analogies Test; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; BWMM = Biemiller Words With Multiple Meanings; BRWI= Biemiller Root Word Inventory WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests. *p<.05; **p \leq .01; ***p \leq .001

Baseline characteristics of final sample of participants (intervention and control)

In order to evaluate the comparability of the intervention group (n=79) and the control group (n=105), the demographic and baseline characteristics of both groups were compared Results of chi-square tests and t-tests (with Bonferronni correction) indicate that there were no statistically significant differences between the control and intervention groups at baseline with respect to age [t(182)=-1.19, p=.24], gender (X^2 =1.91, df= 1, p=.39), home language (X^2 =4.23, df= 1, p=.12), country of birth (X^2 =3.16, df= 1, p=.08), maternal education (X^2 =2.91, df= 2, p=.20), age at arrival in Canada [t(182)=1.78, p=.08], length of residence in Canada

[t(182)=1.88, p=.07], nonverbal ability [t(182)=.96, p=.34], pretest word reading [TOWRE Word Reading: t(182)=-1.68, p=.09; TOWRE Pseudoword Reading: t(182)=-.69, p=.49], vocabulary knowledge [EVT: t(182)=1.69, p=.09; PPVT: t(182)=.32, p=.75], reading comprehension skills [WRMT-R: t(182)=2.15, p=.07; GMRT: t(182)=1.72, p=.09] motivation for reading [t(182)=-1.63, p=.11], and reading strategy use [t(182)=.05, p=.96] (see Table 7 for a summary of pretest characteristics).

Table 7

	Intervention group		Control group)
	(N=79)	(N=79)		
Characteristic (Measure)	n (%)	M (SD)	n (%)	M (SD)
Gender				
Female	51 (64.6)		60 (57.1)	
Male	28 (35.4)		45 (42.9)	
Age		15.03 (1.17)		15.24 (1.20)
Country of Birth				
Canada	23 (57.5)		68 (73.1)	
Outside Canada	17 (42.5)		25 (26.9)	
Age at arrival in Canada		7.53 (5.42)		4.85 (3.59)
Length of residence in Canada		13.53 (3.34)	14.35 (2.44)	
Home language				
English	33 (75.0)		82 (88.2)	
Other	11 (25.0)		11 (11.8)	
Maternal Education				
Secondary level or below	15 (33.3)		40 (42.1)	
Postsecondary level or above	10 (22.2)		25 (26.3)	
Not indicated	20 (44.4)		30 (31.6)	
Nonverbal ability (MAT)		52.02 (10.98)		50.47 (10.82)

Baseline Characteristics of Final Study Sample, by Study Condition
93.00	90.00
82.06 (12.25)	84.95 (11.27)
97.00	99.00
41.54 (12.52)	43.19 (12.06)
87.00	89.00
129.00(24.42)	124.56(18.78)
93.00	90.00
171.96(30.35)	170.55(20.86)
91.00	91.00
46.14(9.91)	43.83(6.30)
93.00	89.00
27.59(10.94)	25.21(9.35)
40.00	42.00
132.89 (25.28)	124.07 (27.32)
3.18 (.74)	3.18 (.66)
	93.00 82.06 (12.25) 97.00 41.54 (12.52) 87.00 129.00(24.42) 93.00 171.96(30.35) 91.00 46.14(9.91) 93.00 27.59(10.94) 40.00 132.89 (25.28) 3.18 (.74)

Note. MAT = Matrix Analogies Test; TOWRE = Test of Word Reading Efficiency; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests; Standard scores based on mean ages at pre- and post-test. Standard scores have a mean of 100 and a standard deviation of 15. T-scores have a mean of 50 and a standard deviation of 10. For all t-tests, p>0.05.

To summarize, the final sample consisted of a total of 184 participants (111 female, 73

male), ages 14 to 20 (M=15.14, SD=1.19). Sixty-eight percent of participants were born in

Canada, while the remaining 32% were born outside of Canada. Eighty-two percent of

participants cited English as one of the languages spoken at home, while 18% indicated that

English was not one of their home languages. Notably, of the entire sample, only 13% of participants came from homes where English was the sole language spoken.

Sample characteristics relative to population normative standards

Participants' pretest standard scores on standardized measures of nonverbal ability, word reading, vocabulary and reading comprehension were used to compare the entire sample of students in this study to their same-age peers with respect to these areas, as findings may be indicative of their relative risk for underachievement compared to normative standards. At the point of program entry, relative to other typically developing youth, on average the sample of participants in this study scored in the Average range on measures of nonverbal ability (MAT: M=93.59, SD= 14.96), word reading (TOWRE Word Reading: M=97.04, *SD*=11.97; TOWRE Pseudoword Reading: M=93.87, *SD*=14.21), vocabulary (EVT-2: M=91.34, *SD*= 16.05; PPVT-4: M=91.09, *SD*=18.08) and reading comprehension (WRMT: 90.37, *SD*=13.54; GMRT: M=41.04, *SD*=18.04). Table 7 summarizes corresponding standard scores for control and intervention groups. These pooled pre-test standard scores were in the Average range for both control and intervention groups, although scores on the GMRT Passage Comprehension task fell within the lower end of the Average range.

2.2.7 Program Fidelity

Tutors were trained to deliver the intervention by doctoral students, as described previously. Program fidelity, defined in this study as the extent to which tutors applied the pedagogical principles designed to contextualize and enhance tutoring effectiveness (i.e., adherence to the program manual), was monitored in the following ways during the current research phase: (a) *A Self-Monitoring Tutoring Checklist*, filled out at the end of each tutoring sessions by the tutors. It included items such as: What was my objective in tutoring today? Did I

manage to achieve this objective? Have I provided rich definitions of new words? Did all the students in the group have opportunities to practice the new words I taught today?; (b) An *Observation-based Monitoring Checklist* which was filled out by trained, graduate level research fellows once a month; (c) The concurrent collection of observational data of every intervention session through audio-recordings of entire sessions across all 13 sessions; (d) For each tutor, audio recordings from 3 randomly selected sessions from early, mid and end points of the program were reviewed and rated by volunteer research assistants followed by a senior graduate student (for reliability) to ensure adherence to manual (i.e., key session components were given a score if completed as per manual, yielding a total score that needed to meet a threshold percentage adherence).

In addition to the self-reported fidelity by tutors, facilitators at each tutoring session were required to check the tutors' fidelity checklists to ensure they had been completed appropriately. To ensure quality of implementation, during the second program session for each cohort, senior level graduate student program coordinators conducted a fidelity check "sit in" with all of the tutors, where they were able to observe program implementation in-vivo. Each tutor received documented personalized feedback, which included information about any program delivery issues that arose during the fidelity check, and how those issues were handled. During the 4th week (session 4), tutors received a collective training refresher via email with the goal of addressing problem areas observed in the 2nd week and/or issues reported by tutors. Overall, results of analyses of program fidelity revealed a mean fidelity of 87% (median = 92%).

2.2.8 Data Analysis

Assessing the Impact of the VLP

To investigate intervention effects on the dependent variables (i.e., measures of vocabulary knowledge and reading comprehension) over time (pre to post) for the intervention group in comparison to the control group, 2 (condition: intervention vs. control) X 2 (time: T1 vs. T2) repeated measures analyses of variance (ANOVA), using a mixed design approach, were conducted. The effect size reported for the overall model for each dependent variable was partial eta squared (η^2). Values less than .06 were considered small effect sizes, values of .06 to .14 medium effect sizes, and values greater than 0.14 large effect sizes (Green & Salkind, 2011). *Assessing the Impact of the VLP on Subgroups of Participants*

Although the overall estimate of intervention effects provides some insight into the intervention's impact at the population level, this approach does not provide any information about the relative efficacy of the intervention (Lesaux et al, 2014). There are likely other dimensions that can influence individual students' intervention needs (Snyder et al, 2017). Thus, it was important to determine the subgroups of participants for whom the intervention was especially effective or less beneficial. To investigate whether the treatment effect varied as a function of student's pretest language comprehension (LC) skills (low LC vs. average and above average LC), pretest levels of reading strategy use (low vs. moderate to high) and pretest levels of reading motivation (high vs. lower levels), the following steps were taken. First, for each area under investigation (e.g., language comprehension), the overall sample was divided into two related subgroups reflecting participants' scores on the relevant measures (e.g., low vs. average motivation). The decision to retain only two categories per variable was made in order to maximize the number of participants per category for analyses.

In order to create two language comprehension subgroups (low vs. average/above average), the sample was divided based on whether students' pretest language comprehension composite scores were either one standard deviation below the sample mean (defined as the "poor language comprehenders", n=23), or were at or above the mean ("good language comprehenders", n= 161). The process of derivation of the language comprehension composite, a summary variable representing the combination of individual pre-test measures of vocabulary and morphological structure, will be described below.

For the two reading strategy use subgroups, the sample was divided based on whether pretest reading strategy levels fell within the low range, represented by a total score of 2.4 or less as defined by the authors of the measure ("poor reading strategy users", n=18), or whether reported reading strategy levels were greater than 2.4 ("good reading strategy users", n=93). To form the two reading motivation subgroups, the sample was divided based on whether students' pretest scores on the measure of motivation to read were either one standard deviation below the mean ("poorly motivated readers", n=17) or were at or above the mean ("well motivated readers", n=95).

To evaluate pretest differences between control and intervention group participants within each subgroup on the relevant outcome measures, independent t-tests were first conducted. Next, in order to evaluate treatment effects, a 2 (time: Time 1 vs. Time 2) x 2 (study condition: intervention vs. control) mixed ANOVA was conducted for each subgroup. For the language comprehension subgroups, treatment effects were assessed on a reading comprehension outcome measure (WRMT-R passage comprehension task). For the remaining subgroups, treatment effects were assessed on both a reading comprehension and vocabulary outcome measure (EVT and WRMT-R). P-values lower than .05 were considered significant. The effect size reported for the overall model for each dependent variable was partial η^2 . As previously described, values less than .06 were considered small effect sizes, values of .06 to .14 medium effect sizes, and values greater than 0.14 large effect sizes (Green & Salkind, 2011). Given the number of subgroup analyses to be performed, the issue of multiplicity was of concern. Specifically, the probability of a false positive finding typically increases with increases in the number of subgroup analyses (Wang & Ware, 2013). A Bonferonni correction was applied to address this issue. This involved adjusting the p-value (α) to ensure that the family-wise error rate was less than or equal to α . Thus the required significance level for the subgroup analyses based on an adjusted α value of .05 was p <.01.

Assessing Predictors of Reading Comprehension Gains among Intervention Participants

In order to examine predictors of change over time in reading comprehension performance, residualized gain scores were computed for a measure of reading comprehension (i.e. WRMT-R). Residualized gain scores examine change as the difference between the observed score at time 1 and the predicted score at time 2 based on the time 1 score (Dalecki & Willits, 1991). Residualized gain scores were used in the present study because they have several advantages over simple difference scores: they are uncorrelated to the time 1 scores and are not influenced by initial scores; they correct for regression to the mean; they are considered more reliable because they are adjusted for unreliability, and they yield more power for studies with a smaller sample size because pre-test predictor scores do not need to be also entered into the model to control for time 1 scores (Dalecki & Willits, 1991). To compute the residualized gain scores, linear regressions were conducted with the independent variable as the time 1 score (pre time point), and then the predicted time 1 score was removed from the time 2 score (post time point). In the model, predictor variables were pretest scores on measures of nonverbal ability (MAT, a control variable), word reading, language comprehension (LC), motivation for reading (MR) as well as age and length of residence (in years) in Canada (control variables). The outcome variable was the residualized gain score for reading comprehension (RES_WPC). The language comprehension predictor variable represents composite scores derived from individual measures of vocabulary and morphological structure (language comprehension composite),

through a factor analysis procedure described below. Prior to a linear regression analysis, a correlation matrix was generated to examine the relationship between the predictors and the outcome variable. Then, a hierarchical linear regression analysis was performed to examine the role of baseline word reading, language comprehension skills and motivation for reading in the prediction of change in reading comprehension performance over time, above and beyond age, non-verbal ability and length of residence in Canada. An overview of all research objectives and corresponding data analyses are presented in Table 8.

Table 8

Overview of Research Objectives and Data Analyses

Re	search	Objective	Da	ta Analysis
1.	Explor to a pr and re	re the effects of reading intervention relative actice-as-usual control group on vocabulary ading comprehension		
	a.	Evaluate and compare intervention effects from pre to post, for control and intervention group on multiple measures of post-test vocabulary and reading comprehension	a.	Two-way mixed analysis of variance (ANOVA)
	b.	Evaluate intervention effects for subgroups of participants	b.	Two-way mixed analysis of variance (ANOVA)
2.	Exami gains (ne predictors of reading comprehension (pre to post) for intervention group		
	a.	Determine relationship between predictors and outcome variable	a.	Correlation analysis
	b.	Evaluate whether (after controlling for age, nonverbal ability and length of residence in Canada), baseline word reading, baseline language comprehension and motivation for reading are predictive of reading comprehension gains	b.	Hierarchical linear regression analysis

2.3 Results

2.3.1 Exploration of the effects of the VLP on youth vocabulary and reading comprehension

In order to examine the change in vocabulary and reading comprehension scores over time, the means and standard deviations of each of the multiple measures of these outcomes were computed at pre and post. As indicated in Table 9, vocabulary and reading comprehension scores across measures increased at post-test, for both intervention and control group participants.

Table 9

	Intervention G	roup (N=79)	Control Group	(N=105)
	Pre	Post	Pre	Post
Outcome Measure	M (SD)	M (SD)	M (SD)	M (SD)
Vocabulary				
Expressive vocabulary (EVT-2)	129.00(24.42)	133.18(24.03)	124.56(18.78)	130.47(18.44)
Standard Score	93.00	94.00	90.00	92.00
Receptive vocabulary (PPVT-4)	171.96(30.35)	177.62(27.79)	170.55(20.86)	174.59(19.63)
Standard Score	91.00	92.00	91.00	92.00
Academic vocabulary (BWMM)	16.90(3.26)	17.14(2.61)	17.03(2.33)	17.02(2.60)
Reading comprehension				
Passage Comprehension (WRMT-R)	46.14(9.91)	46.87(8.80)	43.83(6.30)	45.42(8.41)
Standard Score	93.00	93.00	89.00	93.00
Reading Comprehension (GMRT)	27.59(10.94)	27.88(11.21)	25.21(9.35)	26.06(9.29)

Note. Raw scores were used. Standard scores based on mean ages at pre- and post-test. Standard scores have a mean of 100 and a standard deviation of 15. EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; BWMM = Biemiller Words With Multiple Meanings; WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests

Evaluation of intervention effects from pre to post on outcome variables

Intervention effects on vocabulary tasks. To investigate gains in vocabulary knowledge over time for the intervention and control groups, a 2 (condition: intervention vs. control) X 2(time: pre vs. post) mixed ANOVA was calculated for each of the vocabulary outcome measures (see Table 10 for a summary).

Table 10

Means, Standard Deviations and Results of	of Mixed ANOVAs	for Intervention 1	Effects
/		./	././

Outcome	Outcome	C		Posttest	Group effect/ANOVA
variable	measure	Group	Pretest M(SD)	M(SD)	Group x Time effect
Vocabulary	Expressive vocabulary <i>df</i> (1,182)	Intervention Control	129.00 (24.42) 124.56 (18.78)	133.18 (24.03) 130.47 (18.44)	F=1.85, <i>p</i> =0.18, η ² =0.013
	Receptive vocabulary <i>df</i> (1,182)	Intervention Control	171.96 (30.35) 170.55 (20.86)	177.62(27.79) 174.59(19.63)	F=0.64, p =0.43, η^2 =0.003
	Academic vocabulary <i>df</i> (1,182)	Intervention Control	16.90 (3.26) 17.03 (2.33)	17.14 (2.61) 17.02 (2.60)	F=0.42, <i>p</i> =0.52, η ² =0.002
Reading Comprehension	WRMT-R Passage Comp. <i>df</i> (1,182)	Intervention Control	46.14 (9.91) 43.83 (6.30)	46.87 (8.80) 45.42 (8.41)	F=0.59, p =0.45, η^2 =0.004
	GMRT Reading Comp. <i>df</i> (1,182)	Intervention Control	27.59 (10.94) 25.21(9.35)	27.88 (11.21) 26.06 (9.29)	F=0.24, <i>p</i> =0.63, η ² =0.001

Note. Effect sizes reported are partial eta squared.

Expressive vocabulary. On the EVT-2, there was a significant main effect of time [Wilk's λ =0.69, F (1, 182) = 62.66, p < .001, partial η^2 = .31], indicating that post-test expressive vocabulary scores were significantly higher than pre-test scores on average, with a large effect size as indicated by partial eta squared. The main effect for study condition on expressive vocabulary was not significant, [F(1, 1) = 0.97, p > .05, partial η^2 = 0.007], indicating that intervention and control group scores were not significantly different on this task. Importantly, there was no interaction effect [Wilk's λ =0.99, F(1,182)=1.85, p > .05, partial η^2 = .013], indicating that the vocabulary gains made by the intervention group were not significantly different from gains made by the control group from pre- to post-test.

Receptive vocabulary. On the PPVT-4, there was a significant main effect of time, [Wilk's λ =0.89, F (1, 182) = 22.77, p < .001, partial η^2 = .11], but not for condition, [F (1, 1) = 0.40, p > .05, partial η^2 =0.002]. There was no interaction effect, [Wilk's λ =0.98, F(1,182)=0.64, p > .05, partial η^2 =0.003]. Examination of pre and post-test mean scores on the PPVT-4 indicate that both control and intervention group participants' receptive vocabulary scores significantly increased from pre to post-test and the effect size was in the medium range. Intervention and control group participants' scores were not significantly different, and the difference in mean scores from pre to post-test (i.e., receptive vocabulary gains) between intervention and control group participants was not significant.

Academic vocabulary. On the BWMM, the main effects of time, [Wilk's λ =0.99 F (1, 182) = .34, p > .05], and condition, [F (1, 1) = 0.114, p > .05] were not significant. There was no interaction effect, [Wilk's λ =0.99, F(1,182)=0.42, p > .05]. These results indicate that control and intervention group participant's academic vocabulary scores did not increase significantly from pre to post-test, the scores between study conditions did not significantly differ, and that there

were no significant differences in the mean scores from pre to post-test (i.e., gains in academic vocabulary) between participants in either study condition.

Intervention effects on reading comprehension tasks. In order to evaluate gains in reading comprehension skills over time for the intervention and control groups, a 2 (condition: intervention vs. control) X 2(time: pre vs. post) mixed ANOVA was calculated for both of the reading comprehension outcome measures.

On the passage comprehension task of the WRMT-R, there was a significant main effect of time, [Wilk's λ =0.97, F (1, 182) = 4.28, p < .05, partial η^2 = .03], but not for condition, [F (1, 1) = 1.97, p > .05, partial η^2 =.01]. There was no interaction effect, [Wilk's λ =0.99, F(1,182)=.59, p > .05]. Examination of pre and post-test mean scores on the WRMT-R indicated that both control and intervention group participants' reading comprehension scores on this task significantly increased from pre to post-test, however the effect size was small. Passage comprehension scores of control and intervention group participants did not significantly differ from one another, and the difference in mean scores from pre to post-test (i.e., passage comprehension gains) between intervention and control group participants was not significant.

On the reading comprehension task of the GMRT, the main effects of time, [Wilk's λ =0.99 F (1, 182) = .94, p > .05], and condition, [F (1, 1) = 2.28, p > .05] were not significant. There was no interaction effect, [Wilk's λ =0.99, F(1,182)=0.24, p > .05]. Control and intervention group participant's reading comprehension scores on this measure did not increase significantly from pre to post-test, scores did not significantly differ between the two study conditions, and there were no significant differences in the mean scores from pre to post-test (i.e., reading comprehension gains) between participants in either study condition.

2.3.2 Evaluation of intervention effects for subgroups of participants

Subgroup correlations. Bivariate Phi coefficients indicated non-significant associations of the language comprehension subgroups (i.e., good and poor language comprehenders) with reading strategy use subgroups (i.e., good and poor reading strategy users) (ϕ =.13, p=.19) and with motivation to read subgroups (i.e., motivated and poorly motivated readers) (ϕ =.04, p=.66). Reading strategy use subgroups and motivation to read subgroups were moderately associated (ϕ =.49, p<.001) indicating that categorization of a participant as a good reading strategy user was positively associated with categorization of the student as a motivated reader (and poor reading strategy use categorization positively associated with being a poorly motivated reader).

Subgroup intervention effects. Table 11 presents the results from mixed ANOVAs to investigate differential intervention effects through subgroup analyses. The findings suggest that the effect of intervention significantly varied for a particular subgroup of students.

Table 11

Subgroup	Outcome	Group	Pretest M(SD)	Posttest M(SD)	Group effect/ANOVA Group
Subgroup	variables	Oloup	M(SD)	M(SD)	x Thile effect
Good language comprehenders	WRMT-R <i>df</i> (1,121)	Intervention	47.91 (8.07)	48.75 (6.59)	F=0.34, p =0.29, η^2 =0.003
		Control	43.85 (5.32)	46.61 (7.39)	
Poor language comprehenders	WRMT-R <i>df</i> (1,18)	Intervention	32.35 (8.02)	33.09 (8.08)	F=0.05, p =0.86, η^2 =0.003
		Control	35.95 (6.29)	36.67 (9.64)	
Good reading strategy users	EVT-2 <i>df</i> (1,67)	Intervention	136.07 (18.59)	138.81 (17.06)	F=1.87, <i>p</i> =0.19, η ² =0.03
		Control	126.99 (16.36)	132.95 (18.23)	

Means, Standard Deviations and Results of Mixed ANOVAs for Selected Intervention and Control Participant Subgroups at Pre- and Post-test

WH df(RMT-R (1,68)	Intervention	45.90 (6.44) 44.35 (5.70)	47.24 (6.36) 47.04	F=1.11, <i>p</i> =0.30, η ² =0.02
		Control	(3.79)	(0.00)	
Poor reading EV strategy users df(/T-2 (1,13)	Intervention	130.29 (22.28)	131.43 (20.77)	F=4.21, $p=0.06$, $\eta^2=0.05$
			122.18	130.00	
		Control	(9.45)	(14.06)	
WI	RMT-R		46.47	49.29	
df((1,13)	Intervention	(7.60)	(9.71)	F=0.47, p =0.52, η^2 =0.03
			40.16	39.00	
		Control	(5.04)	(11.19)	
Motivated EV	/T-2		125.11	135.52	
readers <i>df</i> ((1,72)	Intervention	(15.06)	(16.57)	F=8.51, $p=0.01$, $\eta^2=0.11^*$
U V			136.97	139.00	
		Control	(18.94)	(20.33)	
WI	RMT-R		45.27	46.28	
df((1,72)	Intervention	(8.01)	(7.62)	F=0.51, $p=0.48$, $\eta^2=0.01$
U V			43.77	45.78	
		Control	(5.57)	(7.80)	
Poorly					
motivated			111.00	119.00	
readers EV	/T-2 <i>df</i> (1,9)	Intervention	(12.49)	(14.73)	F=2.12, $p=0.18$, $\eta^2=0.19$
	0		129.93	133.63	
		Control	(11.86)	(15.63)	
WI	RMT-R		39.00	40.67	
df((1.9)	Intervention	(6.08)	(3.06)	$F=0.001, p=0.98, n^2=0.00$
uj (41.93	45.00	- 0.001, p 0.00, ij =0.00
		Control	(4.80)	(5.61)	

Note. Effect sizes reported are partial eta squared. $*p \le .01$

Subgroup: Motivated Readers

Independent t-tests indicated that there were no significant differences between the intervention (M=45.27, SD=8.01) and control group (M=43.77, SD=5.57) on passage comprehension scores at pretest, t (93) =1.07, p=0.29. On the EVT-2, the control group scored significantly higher (M= 136.97, SD=18.94) than the intervention group (M=125.11, SD=15.06) at pretest, t (93) =3.36, p=.001.

Intervention effects on vocabulary. Results of mixed ANOVA indicated that on the EVT-2 there was a significant main effect of time, Wilk's λ =0.74, F (1, 72) =25.56, p <.001, partial η^2 = .262. There was also a significant main effect for study condition, F (1, 72) = 4.55, p = .04, partial η^2 = .059, suggesting that the control group scored significantly higher on vocabulary than the intervention group on average. Importantly, there was a significant Condition x Time interaction effect with a medium effect size which indicated that the intervention group made greater gains on vocabulary from pre-test to post-test than the control group, [Wilk's λ =0.89, F(1,72)=8.51, p = .005, partial η^2 = .106] among participants who demonstrated good motivation to read.

Intervention effects on reading comprehension. On the passage comprehension task of the WRMT-R, the main effects of time, [Wilk's λ =0.98, F (1, 72) = 1.63, p=.21, partial η^2 = .022], and study condition, [F (1, 72) = .38, p=.54, partial η^2 = .005] were not significant. There was no significant Condition x Time interaction effect, [Wilk's λ =0.99, F(1,72)=0.51, p=.48, partial η^2 = .007).] This indicates that among motivated readers, control and intervention group participants' reading comprehension scores did not increase significantly from pre to post-test, and there were no significant differences in the mean reading comprehension gains from pre to post-test between participants in either study condition.

Subgroup: Poorly Motivated Readers

Independent t-tests indicated that there were no significant differences between the intervention (M=39.00, SD=6.08) and control group (M=41.93, SD=4.80) on passage comprehension scores at pretest, t(15)=.92, p=0.36. On the EVT-2, the control group scored significantly higher (M= 129.93, SD=11.86) than the intervention group (M=111.00, SD=12.49) at pretest, t(15)=-2.49, p=.01.

Intervention effects on vocabulary. On the expressive vocabulary task, results of mixed ANOVA indicated a significant main effect of time, [Wilk's λ =0.52, F (1, 9) = 8.28, p = .02, partial η^2 = .479], but not for condition, [F (1, 9) = 3.30, p=.10, partial η^2 = .268]. There was no interaction effect, [Wilk's λ =0.81, F (1,9)=2.12, p =.18, partial η^2 = .19]. These results indicate that among students who have poor motivation to read, both control and intervention group participants' vocabulary increased from pre-test to post-test. However, vocabulary gains between intervention and control participants were not significantly different.

Intervention effects on reading comprehension. On the passage comprehension task of the WRMT-R, the main effects of time, [Wilk's λ =0.86, F (1, 9) = 1.43, p=.26, partial η^2 = .137], and study condition, [F (1, 9) =1.61, p=.24, partial η^2 = .152] were not significant. There was no significant Condition x Time interaction effect, [Wilk's λ =1.00, F (1, 9) =0.001, p=.98, partial η^2 = .00]. This indicates that among poorly motivated readers, control and intervention group participants' reading comprehension scores did not increase significantly from pre to post-test, and there were no significant differences in reading comprehension gains from pre to post-test between participants in either study condition.

Subgroup: Good language comprehenders

Independent t-tests conducted to explore group differences on baseline scores at pretest showed that among good language comprehenders, passage comprehension scores on the WRMT-R were significantly higher for the intervention group (M=47.91, *SD*=8.07) than for the control group (M=43.85, SD=5.32) at Time 1, t(141)=3.81, p<.001.

On the passage comprehension task of the WRMT-R, mixed ANOVA revealed a significant main effect of time, [Wilk's λ =0.97, F (1, 121) = 3.60, p=.03, partial η^2 = .03], indicating that both control and intervention group participants' reading comprehension scores on this task significantly increased from pre to post-test, with an effect size in the small range as

indicated by partial eta squared. The main effect for condition was significant, [F (1, 121) = 7.38, p=.01, partial η^2 = .06], indicating that the intervention group had higher reading comprehension scores than the control group. Importantly, the interaction effect, [Wilk's λ =0.99, F(1,121)=.343, p=.288, partial η^2 = .003], was non-significant, indicating that there were no significant differences in the mean scores from pre to post-test (i.e., gains in reading comprehension) between participants in either study condition among good language comprehenders.

Subgroup: Poor language comprehenders

Independent t-tests indicated that there were no significant differences between the intervention (M=32.35, SD=8.02) and control group (M=35.95, SD=6.29) on passage comprehension scores at pretest among poor language comprehenders, t(21)=1.18, p=0.24. Mixed ANOVA results indicated that the main effects of time, [Wilk's $\lambda=0.99$, F (1, 18) = .45, p=.51, partial η^2 = .002], and study condition, [F (1, 18) = 1.19, p=.29, partial η^2 = .062 were not significant]. There was no significant Condition x Time interaction effect, [Wilk's $\lambda=0.99$, F(1,18)=0.05, p=.86, partial η^2 = .003], indicating that there were no significant differences in the mean scores from pre to post-test (i.e., gains in reading comprehension) between participants in either study condition among poor language comprehenders.

Overall, these results indicate that there were no differences in intervention effects among poor and good language comprehenders. Participants in the intervention and control groups who were good language comprehenders at pretest demonstrated significant gains in reading comprehension from pretest to post-test; however participants in the intervention and control groups who were initially poor language comprehenders did not demonstrate significant reading comprehension gains over time. Reading comprehension scores of intervention group participants were significantly higher than control participants' scores among good language comprehenders, and not significantly different among poor language comprehenders. Interaction effects, which account for relative changes in scores (i.e. reading comprehension gains) over time between the study groups (intervention vs control), were non-significant for both subgroups. *Subgroup: Good Reading Strategy Users*

Independent t-tests indicated that there were no significant differences between the intervention (M=45.90, SD=6.44) and control group (M=44.35, SD=5.79) on passage comprehension scores at pretest, t (91) =1.15, p=0.25. On the EVT-2, the intervention group scored significantly higher (M= 136.07, SD=18.59) than the control group (M=126.99, SD=16.36) at pretest, t (91) =2.4, p=.02.

Intervention effects on vocabulary. On the expressive vocabulary task, results of mixed ANOVA indicated a significant main effect of time, [Wilk's λ =0.74, F (1, 67) = 23.33, p < .01, partial η^2 = .258], but not for condition, [F (1, 67) = 2.59, p=.11, partial η^2 = .037]. There was no Condition x Time interaction effect, [Wilk's λ =0.97, F (1, 67) =1.87, p =.19, partial η^2 = .027]. These results indicate that among students who are good reading strategy users, both control and intervention group participants' vocabulary increased from pretest to post-test, with an effect size in the large range as indicated by partial eta squared. Vocabulary gains between intervention and control participants were not significantly different.

Intervention effects on reading comprehension. On the passage comprehension task of the WRMT-R, there was a significant main effect of time, [Wilk's λ =0.94, F (1, 68) = 4.54, p = .04, partial η^2 = .063], but not for condition, [F (1, 68) = .09, p=.77, partial η^2 = .001]. There was no significant Condition x Time interaction effect, [Wilk's λ =0.98, F(1,68)=1.11, p =.30, partial η^2 = .016]. Thus, among good reading strategy users, participants in both the control and intervention group demonstrated increased reading comprehension scores from pretest to post-test with an effect size in the small range. However, mean reading comprehension gains between intervention and control participants were not significantly different.

Subgroup: Poor Reading Strategy Users

Results of independent t-tests indicated that on the passage comprehension task of the WRMT-R, the intervention group scored significantly higher (M= 46.47, SD=7.60) than the control group (M=40.16, SD=5.04) at pretest, t(16)=2.16, p=.03. On the EVT-2, there were no significant pretest score differences between the intervention (M= 130.29, SD=22.28) and control (M=122.18, SD=9.45) groups , t(16)=1.07, p=.28.

Intervention effects on vocabulary. Results of mixed ANOVA indicated that on the EVT-2, there was a significant main effect of time, [Wilk's λ =0.62, F (1, 13) =7.94, p =.02, partial η^2 = .379], but not for condition, [F (1, 13) = 0.26, p = .62, partial η^2 = .019]. The interaction effect approached significance, [Wilk's λ =0.76, F (1, 13)=4.21, p = .06, partial η^2 = .05]. Examination of pre and post-test mean scores on EVT-2 indicate that both control and intervention group participants' vocabulary scores significantly increased from pre- to post-test, with an effect size in the large range. However, mean vocabulary gains for intervention and control participants were not significantly different between the two conditions.

Intervention effects on reading comprehension. On the passage comprehension task of the WRMT-R, there was a significant main effect of study condition, [F (1, 13) = 5.68, p=.04, partial η^2 = .303], but not for time, [Wilk's λ =0.99, F (1, 13) = .15, p = .73, partial η^2 = .011]. There was no interaction effect, [Wilk's λ =0.97, F (1,13)=.47, p =.52, partial η^2 = .034]. Thus, among poor reading strategy users, intervention group participants had significantly higher mean reading comprehension scores than the control group (large effect size); however reading comprehension gains from pre- to post-test were not significantly different between the two study conditions.

Overall, these results indicate that there were no differences in intervention effects between poor and good reading strategy users. Participants in the intervention and control groups who were good reading strategy users at pretest demonstrated significant gains in vocabulary and reading comprehension from pretest to post-test, and participants from both study groups who were initially poor reading strategy users demonstrated significant gains in vocabulary, but not reading comprehension over time. Vocabulary and reading comprehension gains from pretest to post-test were not significantly different between the intervention and control participants, among both poor and good reading strategy users.

Summary of Findings from Subgroup Analyses

In summary, results of the various subgroup analyses indicate that the intervention was effective with respect to improving vocabulary knowledge among participants who were highly motivated readers at pretest. The intervention was not differentially effective for participants who were initially: poorly motivated to read, good or poor language comprehenders and good or poor reading strategy users.

2.3.3 Examination of predictors of reading comprehension gains over time

Derivation of Composite Scores for Selected Predictor Variables

In order to assess the roles of language comprehension and word reading in the prediction of reading comprehension gains post-intervention, it was necessary to utilize predictor measures of language comprehension and word reading skills. However in this study there were multiple theoretically correlated measures of language- and word-based skills administered to participants. Specifically, the measures of language-related skills (i.e., EVT-2, PPVT-4, BWMM, Biemiller Root Word Inventory (BRWI), Carlisle Test of Morphological Structure (CTMS) and word reading skills (i.e., TOWRE Word Reading (TWR), TOWRE Pseudoword Decoding (TPD) all measure different yet related aspects of reading that have been established to be related to reading comprehension. In order to avoid problems with multicollinearity, avoid sample size issues, and to obtain conceptually valid clusters of variables to serve as predictors in the regression equation, principal components analysis with varimax rotation and Kaiser normalization was conducted on those seven measures such that variables comprising extracted factors could be combined into relevant composite variables. Eigenvalues equal to or greater than 1.00 were extracted. With regard to the seven variables used (i.e., expressive vocabulary, receptive vocabulary, academic vocabulary, root word inventory, morphological structure task, word reading and pseudoword reading), orthogonal rotation of the variables yielded 2 factors, accounting for 59.36% and 15.02% of the total variance respectively, resulting in a total of 74.38% of the total variance explained. The factor loadings are presented in Table 12. To enhance the interpretability of the factors, only variables with factor loadings greater than or equal to 0.4 were selected for inclusion in their respective factors.

The first factor was interpreted to represent language comprehension and was comprised of the following variables: EVT (expressive vocabulary), PPVT (receptive vocabulary), BRWI (vocabulary), BWMM (academic vocabulary) and CTMS (morphological structure). The second factor represented baseline word reading and was comprised of the following variables: TWR (word reading) and TPD (reading decoding). Based on the results of the factor analyses, two composite variables (i.e., for language comprehension and word reading) were computed for each participant. The language comprehension and word-reading composites were found to be significantly correlated (r=0.48, p<.01), reflecting the strong relationship between them. Both composites were utilized in the subsequent regression analysis to assess the predictive value of pre-test language and word-based skills on reading comprehension gains following intervention.

Table 12

	Factor		
Predictor	1	2	
Expressive vocabulary (EVT-2)	0.87		
Receptive vocabulary (PPVT-4)	0.86		
Academic vocabulary (BWMM)	0.73		
Vocabulary (BRWI)	0.81		
Morphological Structure Task	0.81		
Word Reading (TWR)		0.87	
Pseudoword Decoding (TPD)		0.87	
Academic vocabulary (BWMM) Vocabulary (BRWI) Morphological Structure Task Word Reading (TWR) Pseudoword Decoding (TPD)	0.73 0.81 0.81	0.87 0.87	

Factor Analysis of Theoretically Correlated Predictor Variables (n=79)

Note. Only factor loadings >.4 are depicted for each factor.

Derivation of Residualized Reading Comprehension Gain Scores

In order to examine predictors of change over time in reading comprehension performance among participants who completed the intervention, residualized gain scores were computed for the primary outcome measure of reading comprehension (i.e., WRMT-R Passage Comprehension subtest). This subtest was selected as the primary reading comprehension outcome measure for regression analyses because it is well standardized and widely used in educational and remedial outcome research (Lovett et al., 2008). Residualized gain scores examine change as the difference between the observed score at time 1 and the predicted score at time 2 based on the time 1 score (Dalecki & Willits, 1991). To compute the residualized gain scores for each intervention group participant, linear regressions were conducted with the independent variable as the time 1 score (pre time point), and then the predicted time 1 score was removed from the time 2 score (post time point).

Relationship between predictors and reading comprehension gains

Predictor variables were conceptually organized in terms of *demographic* factors (i.e., age in months and length in years of residence in Canada), *cognitive* factors (i.e., pretest scores on the MAT, which is a measure of nonverbal ability), *linguistic* factors (i.e., word reading

composite scores, language comprehension composite scores), and psychological factors (i.e.,

motivation for reading) (see Table 13 for a summary of descriptive characteristics).

Table 13

Comprehension Gains	
Predictor	M (SD)
Age (months)	180.33 (14.03)
Length of Residence in Canada	13.53 (3.34)
Nonverbal ability	52.02 (10.98)
Word Reading	124.29 (22.54)
Language comprehension	355.6 (62.62)
Motivation for Reading	132.89 (25.28)

Descriptive Characteristics of Predictors of Reading Comprehension Gains

To examine the relationship between these predictors and the outcome (i.e., reading comprehension residualized gain scores on the WRMT-R passage comprehension task), correlations were computed. Results of the correlation analysis are presented in Table 14.

Predictor variables were weakly correlated with residual reading comprehension gains. Despite the weaker correlations of predictors with reading comprehension gains, they were retained in model due to prior literature indicating their roles in influencing reading comprehension outcomes. Individual predictor variables were weakly correlated with each other, with the exception of word reading and language comprehension which were significantly moderately correlated (r=0.48, p<.01). Age and language comprehension were also significantly negatively correlated (r=-0.28, p<.05). Table 14

Correlations of Demographic (age, years in Canada), Cognitive (nonverbal ability), Linguistic (word reading, language comprehension) and Psychological (motivation) Predictor Variables with Reading Comprehension Residual Gain Scores (n=79)

	1	2	3	4	5	6
1. Reading Comp. Residual						
Gain Score						
2. Age (months)	06					
3. No. Yrs in Canada	.16	12				
4. Non-verbal ability (MAT)	.10	.03	12			
5. T1 Word Reading	.20	13	.06	02		
6. T1 Language Comp.	.22	28*	.24	.22	.48**	
7. Motivation for Reading	.33	32	11	02	.18	.15

Note. 1 = Reading comprehension residual gain score; 2 =Age in years; 3= length of residence in Canada; 4= nonverbal ability score at pre-test on the Matrix Analogies Test (MAT); 5= word reading composite score at pre-test; 6= language comprehension composite score at pre-test; 7= Motivation for reading score at pre-test. *Note.* *p < .05, **p < .01, ***p < .001

Examining Predictors of Reading Comprehension Gains

Given that the role of intervention on reading comprehension gains was evaluated in the previous analysis, it was important to further explore the potential predictors of short-term gains in reading comprehension skills of youth who were exposed to the intervention. To investigate the relative importance of linguistic factors (i.e., word reading, language comprehension) and psychological factors (i.e., motivation for reading) in the prediction of reading comprehension gains in the intervention group (n=79), after controlling for demographic (i.e., age, length of residence in Canada) and cognitive factors (i.e., nonverbal ability), hierarchical linear regression analysis was performed. Demographic and cognitive factors were individually sequentially entered into the model as controls, based on established research indicating their significant relationship to reading comprehension (McGrew & Wendling, 2010; Pasqueralla et al., 2012). Person-level linguistic factors were then entered individually according to their theorized developmental sequence of influence on reading comprehension based on the Simple View of Reading (SVR) (Gough & Tunmer, 1986). Lastly, the person-level psychological factor of

motivation to read was entered into the model to assess its relative predictive value when considered jointly with linguistic factors.

Results for the hierarchical regression examining the relationship between the predictors and reading comprehension residualized gains can be found in Table 15. In the first step, age in months was entered and accounted for 2.4% of the variance in the model, $(R^2 = 0.024, p = .43)$. In step 2, length of residence in Canada was entered and accounted for an additional 10.6% of the variance in the model, (R^2 change= .106, R^2 =.130, p =.09). In step 3, nonverbal ability was entered and accounted for an additional 1.6% of the variance, (R^2 change= .016, R^2 =.147, p =.51). Thus, the control predictor variables together accounted for 14.7 % of the variance. In the fourth step, the word reading composite was entered, explaining an additional 7.5% of the variance in the model, (R^2 change=.075, R^2 =.222, p =.15). In step 5, entering the language comprehension composite significantly explained an additional 12.1% of the variance in the model, (R^2 change=.121, R^2 =.343, p =.04). In the final step, entering the motivation for reading score significantly accounted for an additional 7.7% of the variance in the model, (R^2 change= .077, R^2 =.42, p =.04). Overall, these results indicated that after accounting for participants' age, length of residence in Canada and nonverbal ability, participants who had better developed English language comprehension skills and who had greater motivation to read prior to intervention demonstrated greater gains in reading comprehension after completing a 13-week intervention targeting language and reading comprehension skills.

Table 15

Hierarchical Linear Regression Predicting Reading Comprehension Gains on the WRMT-R Passage Comprehension Task, After Controlling Demographic and Cognitive Factors (n=79)

	В	SE B	β	\mathbf{R}^2	ΔR^2
Demographic Factors					
Step 1					
(constant)	10.57	13.47			
Age	06	.08	15	.02	.02
Step 2					
(constant)	18.91	13.83			
Age	08	.07	21		
No. of Years in Canada	34	.20	33	.13	.11
Cognitive Factors					
Step 3					
(constant)	18.86	13.97			
Age	07	.08	18		
No. of Years in Canada	37	.2	36		
Nonverbal ability	04	.06	14	.15	.02
Linguistic Factors					
Step 4					
(constant)	7.55	15.67			
Age	05	.08	14		
No. of Years in Canada	32	.2	32		
Nonverbal ability	04	.06	14		
Word Reading	.06	.04	.28	.22	.08
Step 5					
(constant)	-1.93	15.56			
Age	02	.07	04		
No. of Years in Canada	33	.19	31		
Nonverbal ability	10	.06	31		
Word Reading	.00	.05	.01		
Language Comprehension	.03	.02	.46*	.34*	.12*
Psychological Factors					
Step 6					
(constant)	-17.91	17.76			
Age	.04	.08	0.09		
No. of Years in Canada	31	.18	-0.31		
Nonverbal ability	11	.06	-0.36		
Word Reading	00	.04	-0.01		
Language Comprehension	.04	.02	.50*		
Motivation for Reading	.05	.03	.31*	.42*	.08*

2.4 Discussion

This study addressed the differential impact of the VLP intervention on the vocabulary and reading comprehension outcomes of high school youth, and identified predictors of reading comprehension development among intervention group participants. What follows is a discussion of the specific findings and their implications.

Intervention effects on vocabulary and reading comprehension outcomes

Results from this study indicated that both control and intervention group participants' expressive and receptive vocabulary scores significantly increased from pre- to post-test, with large and medium range effect sizes, respectively. Relative to the business as usual control group, the intervention group did not make significantly more gains across measures of vocabulary. With respect to reading comprehension, results indicated that both control and intervention group participants' reading comprehension scores significantly increased from pre to post-test, with a small effect size, on the primary reading comprehension outcome measure (i.e., WRMT-R passage comprehension subtest). However, average reading comprehension gains of the intervention group did not differ significantly from those of control participants on this measure. The finding that participants in both intervention and control groups exhibited similar improvement across outcome measures is typically an indication that there was no intervention effect. It is possible that the intervention was not sufficiently intensive to yield significant change immediately post-intervention. Fuchs, Fuchs and Vaughn (2014) argue that poor response to research-based intervention can be indicative of a need for more intensive instruction. The fact that vocabulary and reading comprehension scores steadily increased over time is suggestive of skill progression that was perhaps not fast enough. As such, participants may benefit from intensification of instruction which can accelerate learning, by making select modifications to the

intervention (e.g., employing smaller group sizes, increasing frequency of sessions, focusing on different instructional aspects, using reading intervention specialists such as special education staff to administer the intervention rather than volunteer tutors) (Fuchs Fuchs & Vaughn 2014; Fien et al., 2018). Other reading intervention studies have suggested that reading comprehension is less readily malleable to rapid change among older students (e.g., Miciak et al., 2018). Miciak et al. (2018) found null effects on primary reading comprehension outcomes relative to a business as usual control group in their randomized control trial of multicomponent tutoring for fourth graders, but found significant effects on measures of reading fluency. They argued that researchers conducting interventions targeting reading comprehension should expect gradual growth over time in this skill area. It is also possible that although there may not have been sufficient time to achieve immediate improvement in reading comprehension skills, there may have been immediate intervention effects on the more proximal skill of reading strategy acquisition. Given this study's focus on evaluation of more distal program outcomes (e.g., reading comprehension skills following targeted strategy instruction), intermediate skills such as reading strategy acquisition were not measured as direct outcomes in this study, and represents a limitation that should be explored in future studies.

Although the present study's finding of null intervention effects is consistent with similar intervention studies that found minimal to null effects (Therrien & Cook, 2018; Fuchs et al., 2018; Miciak et al., 2018; Fien et al., 2018), it is important to recognize several methodological factors that may have affected the study's ability to demonstrate an effect (Becker et al., 2003) prior to discounting the efficacy of this intervention. In fact, investigation of null effects can have significant implications for educational practice and policy (Therrien & Cook, 2018). The current study results prompt an active exploration and negotiation of such factors that potentially impeded the detection of intervention effects, reflecting an important strength of this study in

advancing knowledge of how best to design community-based interventions in culturally and linguistically diverse contexts.

Wilson and Lipsey (2014), among other researchers, have noted that various implementation problems may result in smaller effects. One consideration is that control participants may have modified their personal behaviors that are related to reading outcomes (e.g., actively looking up unfamiliar vocabulary words, reading more academic texts) simply by virtue of knowing that they were participating in a research project; this is called the Hawthorne or placebo effect (Rossi, Freeman & Lipsey, 1999). There might have also been a testing or practice effect, which occurs when participants improve post-test scores because they learned correct responses to items on outcome measures at pre-test (Becker et al., 2003). The WRMT-R reading comprehension and expressive vocabulary measures were more susceptible to this effect as the same version was administered at post-test, unlike the secondary reading comprehension measure (GMRT), for which a different version of the test form was used for post-testing.

Contamination of the control group was another factor that may have affected the ability to demonstrate an intervention effect. For example, some participants in the control group may have been exposed to part or all of intervention and as a result demonstrated change in outcome measures in a manner similar to those actually receiving the intervention. Sources of contamination included holding sessions in small groups in the same physical environment where other youth in the Pathways to Education program (i.e., the broader community-wide intervention program) attended afterschool homework help and academic tutoring. Control participants in this study had a very high likelihood of being present during intervention sessions (e.g., sitting at a neighboring table, interacting briefly with peers involved in the intervention group). They may have even discussed aspects of the sessions with their intervention group peers. In these ways, control participants may have obtained sufficient information to influence them indirectly in making changes to the application of their own skills, potentially enabling them to achieve outcomes similar to those of intervention group (Becker et al., 2003).

On a related note, treatment effects are typically more likely to be found and be of larger effect when the comparison group receives no services at all or is at least not exposed to the intervention when engaged in a 'business as-usual' experimental condition. Perhaps then, another source of diminished evidence of intervention effects was the fact that the reading intervention was embedded within an already existing broader community intervention providing generalized academic and other supports for students, all while these students continued to receive regular school-based instruction. All of the described effects are even more critical to consider for an intervention involving immigrant and language minority youth in the Canadian context, a context where the youth's families are on the whole more educated based on Canadian immigration policies, especially compared to other countries (Statistics Canada, 2011; OECD, 2018b). Highly educated parents are more likely to encourage their children to make the most of their (multiple) learning contexts and available interventions as they place a higher value on academic success (e.g., Roundfield, Sanchez & McMahon, 2018).

Maturation effects may have also been at play to confound intervention effects (Cook & Campbell, 1979), especially given that the intervention took place over the course of much of an academic term, where all students may have simultaneously received similar instruction at school or experienced developmental changes that increased their general knowledge.

Finally, an aspect of analysis that may have influenced results was the designation of intervention effectiveness based on a comparison of change scores at a single time point immediately following intervention rather than over an extended period or multiple post-intervention time points, as this approach did not allow for consideration of participants' longitudinal trajectories of growth as a result of the intervention. That is, intervention group

participants may have made significantly more gains over time than the control group in vocabulary and reading comprehension that may not have been captured by a single immediate post-test score. As such, a longitudinal research design with multiple measurement time points would have facilitated a comparison of clusters of intervention effects on the basis of growth patterns over time.

Differential intervention effects for subgroups of participants

In a study such as this one, involving complex, multi-determined outcomes, research has indicated that it is important to recognize that the intervention alone may represent only one of multiple potential factors affecting outcomes (Becker et al., 2003). For this reason it was important to conduct subgroup analyses as a next step in understanding intervention effects. Although overall results of the intervention showed little difference in vocabulary and reading comprehension gains between control and intervention groups, the intervention nevertheless helped a subset of participants to improve their vocabulary outcomes, allowing for a consideration of potential moderators of intervention effects. The findings of this study indicate that although overall intervention effects were not significant, there were significant intervention effects for a particular subgroup of participants, namely those who were motivated readers prior to commencing the intervention. Specifically, results of subgroup analyses indicated that when compared to controls, there were no significant differential treatment effects for intervention group participants who were poor language comprehenders, good language comprehenders, good reading strategy users, poor reading strategy users or poorly motivated readers. However, wellmotivated readers in the intervention group demonstrated significant vocabulary gains compared to motivated students who were in the control group, even though students in the intervention group scored significantly lower than control participants on the vocabulary measure prior to intervention. The effect size was in the moderate range. It is possible that the intervention was

sufficiently intensive to yield immediate vocabulary gains among the more struggling yet highly motivated readers. Given this subgroup's lower pretest scores relative to control participants, it is also possible that they self-selected into the intervention group. Overall, the results indicate that motivation to read could be considered a moderator of intervention effects. This finding suggests that at least among those who were highly motivated at pre-test, an intervention that incorporated program elements such as contextually based, direct and rich vocabulary strategy instruction, repeated practice, multiple exposures to words and small group discussion, for example, can be effective in improving vocabulary. This is supported by the existing evidence base with respect to the role of these intervention components in supporting vocabulary development (e.g., Townsend & Collins, 2009; Biemiller & Boote, 2006; Hwang et al., 2015), although the current study did not examine the particular components that may have directly supported vocabulary growth. Other studies have shown differential intervention effects based on differing preintervention factors such as vocabulary (Lesaux et al., 2014; Townsend & Collins, 2009) but fewer studies have implicated baseline reading motivation in differentiating intervention effects specifically (e.g., Guthrie, McCrae & Lutz Klauda, 2007), and even more so among culturally and linguistically diverse adolescents. The existing literature does however suggest that motivation to read can be an important factor associated with overall reading and academic achievement, typically reported among younger readers (e.g., Klauda & Guthrie, 2015; Logan, Medford & Hughes, 2011; Stutz, Schaffner & Schiefele, 2016) and middle school students (e.g., Lesaux et al., 2012). Its role in differentiating intervention effects, particularly among highschool aged adolescents is worthy of further examination.

Predictors of Reading Comprehension Gains

Regression analyses allowed for the joint consideration of the relative contributions of person-level linguistic factors (word reading, language comprehension) and a person-level

psychological factor (motivation to read) to predicting reading comprehension gains. Overall, the findings from regression analyses indicate that after accounting for demographic factors (participants' age, length of residence in Canada) and cognitive factors (nonverbal ability), participants who had better developed English language comprehension skills and who had greater motivation to read prior to the intervention demonstrated greater gains in reading comprehension after completing a 13-week intervention targeting language and reading comprehension skills. Word reading was not a significant predictor of post-intervention reading comprehension gains; it did not contribute to the variance in outcome scores in this study's model. The finding that language comprehension is a significant predictor is consistent with the SVR model (Gough & Tunmer, 1986; Hoover & Gough, 1990) which in part postulates that clusters of variables associated with language comprehension form one of two main interacting elements that underlie reading comprehension (Geva & Wiener, 2015), with the other element being word-level reading skills. Prior research has indicated that among second language learners, word reading skills remain significantly predictive of reading comprehension even in adolescence for those whose exposure to the second language initiates in adolescence (Pasquarella et al., 2012); thus age of initial exposure to a second language is an important consideration (Geva & Wiener, 2015). In the present study, the contribution of word reading skills to the variance in reading comprehension gains may have been attenuated by virtue of controlling for participants' length of residence in Canada in the model, assuming that duration of Canadian residence serves as a sufficient proxy for duration of exposure to English. That is, word reading skills are not predictive of reading comprehension gains once duration of exposure to English has been considered, particularly in this sample that consisted of a greater proportion of Canadian-born students. Motivation, including motivation to read, has been shown by previous research to be associated with reading comprehension outcomes (e.g., Klauda &

Guthrie, 2015; Stutz et al., 2016) and is also predictive of improved reading and academic achievement overall, consistent with the findings of this study. In summary, the combination of analyses in the present study suggests that the person-level factors of motivation to read and language comprehension are implicated in the differentiation of program effects.

Study Limitations and Future Directions

Methodological considerations. It is important to acknowledge that findings from subgroup analyses were limited by the criteria (i.e., cut-offs) used for defining subgroup membership. The application of such cut-offs on constructs that are inherently dimensional rather than categorical (i.e., motivation to read) is problematic, as it results in a loss of potentially valuable information about dimensional aspects not captured by discrete categories. Criteria used for the various subgroup categorizations split the sample into low (below the mean) and high (at or above the mean) subgroups per variable (i.e., self-reported motivation to read, strategy use, language comprehension score). Although delineation of only two categories per variable was employed to maximize subgroup sample sizes for comparisons, this approach resulted in unbalanced subgroup sample sizes per variable, with subgroups comprised of participants with lower scores having lower sample sizes. Evaluation of effect sizes in instances of nonstatistically significant findings can be instrumental in determining whether a larger sample size may have resulted in statistical significance. Medium to large effect sizes suggest that differences in mean reading comprehension or vocabulary gains between control and intervention group participants would have reached statistical significance with a larger sample size. Results from this study indicated a medium effect size of the Condition x Time interaction among poorly motivated readers(η^2 =.19), suggesting that with a larger sample size, the results would demonstrate that poorly motivated readers who receive the VLP intervention make significantly

more gains in expressive vocabulary than poorly motivated readers in the business as usual control group. Other non-statistically significant findings yielded low effect sizes.

It is also possible that differential treatment effects may have been observed if subgroup outcomes were compared at follow-up time points beyond the two time points employed in this study, particularly if environmental influences were potentially at play in influencing outcomes (Becker et al 2003). This study would benefit from replication with a larger sample to allow for analyses such as latent class growth analysis (LCGA) and growth mixture modeling (GMM) (Ram & Grimm, 2009) which could be used instead of hierarchical linear regression. These approaches support the process of identifying homogenous subpopulations within the larger heterogenous population of students, for enhanced identification of meaningful groups and interpretation of differential effects (Jung & Wickrama, 2008). Findings from the regression analysis are also limited by the specific predictors used in the model, as different predictive models may yield different results (Field, 2009). However, the current prediction model reveals the importance of jointly considering the role of linguistic and psychological person-level factors in predicting reading comprehension gains.

The sampling approach of this study did not allow for random assignment of participants to the intervention or control groups. In fact, the comparison group may not be reflective of a true control sample given the possibility that participants self-selected into the intervention based on their level of interest or willingness to participate. As well, some of the participants assigned to the control group had in fact completed a minimal number of intervention sessions (e.g., up to 2 sessions) prior to being transferred to the control group due to intervention attrition or elective withdrawal from participation in the intervention. The approach of retaining such participants was taken in order to maximize the comparison group sample size and hence subsequent analyses. However, this may have contributed to comparison group contamination and may have

minimized the comparative value of the control group based on unexpected and unaccounted for pre-existing differences (e.g., relatively less motivated participants).

In order to improve the likelihood of detecting intervention effects if present, future research would benefit from careful application of approaches to minimize threats to randomization (e.g., differential self-selection into control or intervention groups). For example, this may be achieved through increased efforts at intentional recruitment and retention of students with diverse baseline levels of interest and motivation to participate in the overall study. To reduce the contamination of control group participants by intervention group participants, measures can be taken to create more physical distance between the study conditions such as holding intervention sessions in a separate and enclosed environment to which control participants do not have access. This way, opportunities for inadvertently exposing comparison group participants to aspects of the intervention are minimized (Becker et al., 2003). Additionally, the use of measures of known confounding variables would facilitate account for their potential moderating impact in the analyses (e.g., a measure that tracks the amount of intervention sessions to which controls were unintentionally exposed). Finally, the use of a longitudinal methodological approach would be facilitative of making more definitive inferences about the true nature of the relationships under study. The use of multiple time points in this study for example, may have allowed for the comparison of the rate of skill growth over time between control and intervention group participants.

Study Implications

In the Canadian context, this study represents the first quasi-experimental evaluation of a research-based after school intervention targeting the vocabulary and reading comprehension skills of culturally and linguistically diverse high school-aged youth. This study's evaluation of differential effects based on pre-intervention student characteristics allows for the identification

of important pre-intervention participant characteristics associated with differential outcomes, and informs the design of interventions to better address these factors. Thus, it highlights the utility of determining which students to target for similar reading interventions to be most effective, or the importance of making decisions about which intervention aspects require modification so that more students can benefit (Biemiller & Shany, 2010). For example, the VLP could be differentiated to address varying reading motivational levels. As previously noted, in this intervention evaluation study involving complex outcomes with multiple causes, it is important to recognize that investigations of general intervention efficacy alone limits the understanding of the multiple potential factors affecting outcomes (Becker et al., 2003). Further research endeavors that utilize mixed methodology to examine other person-level as well as broader contextual factors impacting response to intervention, particularly given the heterogeneity of youth and their responses to intervention, are essential.
Chapter 3: An Exploration of Contextual Factors Associated with Response to Reading Intervention with Marginalized Adolescents (Study 2)

3.1 Introduction

Over the past decade, there has been a surge in research efforts focused on how best to enhance reading skills in adolescents, particularly immigrant, linguistic minority and marginalized students (e.g., Ebert et al., 2014; Lesaux et al., 2014; Pasquarella et al, 2012; Vadasy & Sanders, 2010). It has been well established that the development of sound reading comprehension ability relies on several cognitive and linguistic skills such as effective phonological processing, phonological short-term memory, reading fluency, language comprehension and so on (e.g., Gough & Tunmer, 1986; Cain et al., 2004; Kendou, Savage & van den Broek, 2009). It has also been well established that reading in a second language involves the same factors and processes at the person-level (e.g., Geva & Farnia, 2012; Gottardo & Mueller, 2009; Verhoeven & van Leeuwe, 2012). What has been garnered from the research in reading skill development over the last few decades is that reading skill development is multidetermined, influenced by many factors beyond the individual learner's personal abilities and characteristics (Geva & Wiener, 2015). This is particularly pertinent to language minority, immigrant and socio-economically disadvantaged youth, who can face complex challenges in the process of seeking positive learning outcomes, as they often have to simultaneously overcome other adversities such as their socio-economic disadvantage, acculturating to a new country, negotiating new personal identities, and language barriers to name a few [Organisation for Economic Cooperation and Development (OECD), 2018b]. Several researchers have noted the importance of acknowledging contextual factors relevant for understanding reading achievement in second language learners (Geva & Wiener, 2015; OECD, 2018b). Although there have been

97

more recent intervention efforts targeting the reading skill development of culturally and linguistically diverse youth, it is imperative that such intervention programs address factors beyond individual-level cognitive and linguistic skills in order to maximize student benefits. *Understanding contextual influences on reading achievement through an ecological framework*

Bronfenbrenner's (1979, 2005) ecological theory of human development provides a valuable framework for considering the multiple influences on reading achievement of language minority and disadvantaged youth, and is the framework that guided the data collection and analysis in this study. According to this theory, human development is shaped over time through interactions between an individual and their broader environment – an environment that is characterized by several nested system levels ranging from immediate individual contexts up to macro-level societal contexts. More specifically, Bronfenbrenner (2005) theorized that an individual's environment consists of the following nested, concentric levels: the microsystem, mesosystem, exosystem and macrosystem.

Microsystem. This system represents the innermost, immediate setting of a person's life. It includes an individual's physical and mental characteristics, personality, temperament and belief systems. The microsystem also includes the individual's reciprocal interactions and any settings in which the individual exists directly (home, school, peer environment, neighbourhood). An individual can interact with his or her immediate setting by engaging in purposeful, goal driven activities such as reading a book or engaging in conversation with someone. As microsystems include other individuals with their own set of characteristics, personality, belief systems and so on, interaction with others through activity (e.g., completing work with the help of a parent or tutor) allows for others in this immediate setting to influence an individual's development. *Mesosystem.* This system is characterized by an individual involved in activities in multiple settings (e.g., home and school). It essentially represents the links between microsystems (e.g., home-school, home-neighbourhood). These different settings can be linked in several ways, such as through relationships across settings (e.g., a student whose sibling attends the same school) or communication across settings (e.g., a school social worker sending a note home to a parent). It is hypothesized that development at this level is dependent on mutually positive, bidirectional interactions between settings (Wilson, 2012).

Exosystem. The exosystem comprises settings in which events that impact the individual (and consequently the individual's development) occur despite the fact that the individual is not an active participant within those settings. An example is a set of policymakers from a different jurisdiction making decisions about educational programs for students residing within a particular neighborhood. Although the students are not direct participants in the microsystem involving the policymakers, the decisions made can have a direct impact on these students. Other examples of an exosystem include an event occurring at a parent's job that impacts their child (or vice versa), or a string of violent acts committed by a criminal group in proximity to a child's neighborhood, resulting in imposed curfews that impact the child's activities.

Macrosystem. Macrosystems are characterized by their constituent meso- and exosystems as well as associated values, ideologies or belief systems that guide activities within the system, such as established laws, standards and norms for individuals' behavior. Some examples of macrosystems are cultures, neighborhoods and educational systems.

Although Bronfenbrenner's ecological systems theory has largely been applied in the field of educational psychology, it can be useful for understanding youth as quite complex social beings who continually shape and are shaped by their immediate and broader environments. It is thus reasonable to extend this theory to help inform the range of complex contextual factors that

can influence reading achievement of language minority and disadvantaged youth. In the case of the present study, this theory will be used as a framework to understand the contextual factors that potentially differentiate students who favorably responded to a reading intervention and those who did not. The theoretical assumption in espousing this approach for the present study is that adolescents' contexts (when explored in relation to their potential impact on academic achievement), as well as the *interaction* between adolescents and their contexts, is suggestive of the likelihood of positive response to targeted academic intervention. The aim then, is to better identify critical contextual aspects that may be targeted to maximize adolescent benefit from interventions.

Individual differences in resilience

The ecological framework is also consistent with the concepts of risk and resilience (Masten, 2004), in that aspects of youth's contexts can function to support achievement and minimize risk, thereby promoting resilience. Research has shown that on the whole, individuals' response to adversity differs when contextual factors are considered (Zolli & Healy, 2012). The concept of resilience has been increasingly utilized by researchers "to identify when, how and why people who have been exposed to negative experiences display less vulnerability" (OECD, 2018b, p. 31). When applied to the field of education, the aim of such research is to explore factors related to the ability of some students to succeed academically despite the disadvantages or risk factors they face, and to understand why certain students' outcomes differ from their peers even when mutually exposed to similar forms of disadvantage (OECD, 2018b; Masten, Powell & Luthar, 2003; Rutter, 2006). With regard to differential response to academic intervention, understanding those factors that facilitate or hinder individuals' attainment of positive intervention outcomes can be a useful step towards identifying practices that promote and support student's positive response to intervention efforts in the face of multiple adversities

(e.g. language and/or ethnic minority status, marginalization, economic disadvantage, adverse neighborhood). Prior research has elucidated the multilevel nature of risk and protective factors that essentially serve to moderate the impact of adversity while supporting academic resilience. These multilevel contextual factors, consistent with Bronfenbrennner's (2005) ecological model, range from individual attributes of an individual student to family characteristics and extend further to aspects of the broader social environment including peers, school, neighborhood and wider community (Masten & Garmezy, 1985; Werner & Smith, 1992; Rutter, 2000; Masten, 2001; Luthar & Cicchetti, 2000).

'Expanded' Simple View of Reading as a Model for Exploring Contextual Differences in Response to Intervention

Developmental research over the last two decades has significantly contributed to an understanding of the factors associated with typical and atypical second language and literacy development of children and adolescents. More specifically, much research in this area has been focused on the development of reading comprehension skills, given that the primary goal of reading is to obtain meaning from text (Geva & Wiener, 2015; Chall, 1996). A variety of models have been suggested to help understand individual differences in reading comprehension ability. One prevalent theory for which there is strong support in the literature as an explanation of individual differences in reading comprehension among first-language learners is called the Simple View of Reading (SVR) (Gough & Tunmer, 1986; Hoover & Gough, 1990). According to the SVR, reading comprehension development is the product of having accurate and fluent word level skills (e.g., decoding pseudowords, spelling, reading words in isolation) and language comprehension skills (e.g., vocabulary, morphology). This model has been well-supported among child and adolescent learners in their first language as well as among students learning English as a second language. The SVR, however, has been critiqued for its intensive focus on

factors contributing to reading comprehension development that are intrinsic to the child or adolescent, such as specific cognitive or linguistic processing skills. Although these factors are critical to examine, Geva & Wiener (2015) expanded the SVR model to account for research on socioemotional, contextual, home, family and sociocultural factors that also play a role in second language literacy and language development. This 'expanded Simple View of Reading' model is consistent with Bronfenbrenner's (2005) ecological model. It espouses that beyond cognitive and linguistic factors accounting for reading development, it is important to explore broader contextual factors such as home (e.g., parental education, family values), school and neighborhood demographic characteristics, as well as relevant policies that address factors such as curricula, capacity building, mental health resources and teacher training (Geva & Wiener, 2015). The model is thus useful for conceptualizing factors contributing to risk and resilience that are representative of sources of individual difference in response to intervention. In the current study, this expanded SVR model was used as a framework for exploring individual, immediate context, family, peer, school, community and cultural factors that differentiated adolescents who demonstrated improved reading outcomes post-intervention compared to those whose outcomes remained unchanged following the same intervention.

Individual-level factors. Multilevel perspectives on risk and resilience (OECD, 2018b) have noted the following factors as particularly relevant contributors to academic achievement as a whole among language minority and socio-economically disadvantaged youth: gender, country of origin, first language oral language proficiency, vocabulary and reading comprehension skills, number of years of residence in the country of immigration, motivation to achieve, engagement in academic-related tasks (Guthrie, 2004), and academic self-concept. With respect to reading outcomes, individual reading behaviors have also been associated with student outcomes. Stanovich (1986) for example noted that reading behavior history (e.g., amount of reading

practice), history of success, failure and reward in the context of academic tasks and motivational differences differentiates good and poor readers.

Immediate contextual factors. Often in order to evaluate the impact of educational programs or interventions, it is important to consider the immediate program context. Factors in the immediate context are comprised of those factors associated with the actual intervention itself, such as specific program content and activities, instructional approach and tools, and so on. These variables have been shown in various learning contexts, particularly within schools, to be critical aspects that can impact the attainment of positive outcomes (e.g., Rubinstein-Avila, 2003; Lesaux et al., 2012) and require examination in order to determine factors that differentially impact response to programs and interventions.

Family context. At the family level, it has been shown that the following factors are associated with learning outcomes: parents' educational attainment, language spoken at home, degree of parental support and involvement with academic life and with the school community (e.g., OECD, 2018b, Hart & Risley,1995; Anisef et al., 2010). With respect to reading achievement specifically, family literacy and engagement in literacy related activities in the home environment have been linked to students' academic achievement (e.g., McKool, 2007; Hart & Risley, 1995; Neuman & Dickinson 2003; Snow et al 1998; Geva & Wiener, 2015). Prior research has also indicated that access to print at home promotes reading comprehension (e.g., Cummins, 2012). Additionally, a relationship has been indicated to exist between students' proficiency in their home language and academic proficiency in the language of instruction among second language learners (e.g., Cummins, 1991; Geva & Ryan, 1993).

Peer context. At the peer contextual level, factors such as peer social support, peer friendships and peer engagement in learning are important areas to consider when examining factors that influence students' learning and reading outcomes in particular. It has been widely

noted that friends and peer groups have a direct influence on adolescent students' academic achievement (e.g., Brown, 1990; Steinberg, Dornbusch & Brown, 1992; Ryan, 2000) especially among ethnic minority students (e.g., Goodenow & Grady, 1993).

School factors. At the school level, the following factors have been described as relevant to consider in understanding student's learning outcomes: school policies, learning environment, school climate, and a sense of belonging or engagement with the school community (e.g., Stanovich & Jordan, 1998; Portes & MacLeod, 1996; Juvonen, 2006).

Community and cultural factors. At the broader community and cultural level, the following factors are also important to consider: neighborhood climate (e.g., degree of violence, cohesiveness, sense of belonging), sensitivity to diversity, available programs and resources to support students/resource allocation, cultural identity, acculturation, systemic discrimination, and educational policies (Ogbu, 1991; Gibson & Ogbu, 1991; Cummins, 2012).

Qualitative Methodology and Program Evaluation

Qualitative methodology has been recognized as a beneficial supplement to intervention outcome research for understanding participants' experiences of intervention and to help in the identification of potential processes or mechanisms associated with change (Yin, 2003; Stake, 2005). Qualitative research can inform quantitative research by allowing for an examination of: 1) who change occurs for, 2) why change takes place for certain participants and not others, and 3) factors that are potentially associated with change (Creswell, 2011). This approach is particularly relevant in multicultural contexts, as qualitative data collection approaches can help to honor the voices or experiences of individual students and shed light on the complexity of reading skill development under circumstances where issues such as language minority status, marginalization, cultural identity and interactions between self and the community are relevant to address. These kinds of contextual factors do not easily lend themselves to quantitative examination or experimental manipulation (Creswell & Plano Clark, 2011). The ability to take advantage of more diverse forms of data collection (e.g., interviews with open ended response formats) provides students the opportunity to elaborate on the issues and circumstances affecting them. In this way, qualitative data and results can help to build on understandings gained from primary outcome research (Creswell & Plano Clark, 2011).

Rationale for the present study

Using an ecological framework as a guide for inquiry, the goal of the current study was to explore the factors differentiating adolescents whose reading comprehension outcomes improved or did not improve following an after-school reading intervention (the Vocabulary Learning Project, or VLP). This study aimed to build on the findings from the prior impact study (Study 1) to help inform the multiple contexts that are potentially associated with differential response to intervention using qualitative methodology.

3.1.2 Research Objectives

The present qualitative study followed from a preliminary quantitative research phase (Study 1). In Study 1, the overall effectiveness of the VLP reading intervention was evaluated (see Study 1 for a description of the intervention). Quantitative analyses during the first phase examined person-level predictors of change as well as the effectiveness of the intervention as a whole in improving vocabulary and reading comprehension outcomes of language minority and socio-economically disadvantaged youth. Findings from that study indicated that intervention group participants did not make more gains post-intervention than business as usual control participants with respect to vocabulary and reading comprehension. However subgroup analyses revealed that those intervention group participants who were highly motivated to improve their reading skills at the outset of the intervention did make more gains than control participants on a

measure of vocabulary. In light of these findings, the current study was conducted as a result of a desire to better understand the factors that differentiate program improvers from non-improvers, and to identify contextual factors not captured by quantitative examination that may be associated with reading skill development. Several researchers have noted that quantitative research approaches in themselves are ineffective for studying the relationship between multiple contextual areas and observed outcomes, such as the outcomes of an intervention (Creswell & Plano Clark, 2011). Thus, more specifically this study aimed to explore the factors at different contextual levels that potentially differentiate adolescents' response to the 13-week reading intervention targeting vocabulary and reading comprehension skills. Contextual areas that framed the data gathering process and hence were explored in this study included the individual level context, immediate intervention-based context, peer, family, community (including school and neighborhood) and cultural contextual levels. Through exploration of relevant factors at each of these contextual levels, the study endeavored to identify the factors that promote intervention gains and factors that do not support positive gains, to ultimately shed light on factors that may be associated with resilience in the face of academic risk. The following research questions were explored in this study:

1) What are participant accounts of person-level factors (motivation, academic selfconcept), peer and family-level factors (e.g., parent involvement, peer support), community and cultural factors (e.g. neighborhood violence, cultural identity) as well as intervention-level factors (e.g., perceptions of tutor, tutoring group, intervention activities and materials) and the impact of these factors on their experience of the program? 2) Which factors differentiate the experiences of program improvers and program non-improvers after accounting for participants' age, grade and nonverbal ability?

3.2 Method

3.2.1 Methodological Approach: Multiple Case Analysis

To address the research questions in this study, a multiple case analysis procedure was employed. The case study approach has been established by prominent qualitative methodologists such as Robert Yin (2003) to be particularly useful in the following circumstances: (a) when answering "why" and "how " research questions (e.g., "why" is an intervention ineffective for certain participants?); (b) when researchers wish to study contextual conditions relevant to the phenomenon under investigation; (c) when the behavior of participants cannot be manipulated, and (d) when there are unclear boundaries between the phenomenon and its context. In fact, Yin (2003) defines a case as "a contemporary phenomenon within its real life context, especially when the boundaries between a phenomenon and context are not clear and the researcher has little control over the phenomenon and context" (p. 13). In the case of this study, if the phenomenon under investigation is defined as "the impact of the VLP reading intervention on an individual student," then the case or item of analysis is the individual student participant. In turn, the primary research question can be framed as "Why do some individual students" respond positively to the VLP reading intervention and others not?" or "How does the individual and their context influence response to intervention?" As previously described, when Yin's (2003) definition of a case is considered in light of Bronfenbrenner's ecological model, which espouses that individuals are embedded within a broader context, it becomes apparent that a case study approach lends itself to exploration of relevant contextual factors influencing individual students' outcomes. In this way, many prominent researchers have noted that case study analytic

approaches are particularly useful for program evaluation (Yin, 2003; Merriam, 1998, Stake, 2005). Similarly, it has been argued that case study analysis is useful when an initial study yields unexpected outcomes and the researcher wishes to address the question of "what happened?" in the initial study phase.

Multiple case studies are one type of case study design, and allow for exploration of differences within and between cases (Baxter & Jack 2008). Multiple case design involves the extensive study of a number of carefully, purposively selected cases to gain a better understanding of a problem and as described above, to theorize about broader contextual factors (Chmiliar, 2010). This methodology involves the analysis and synthesis of the similarities, differences and patterns across multiple cases (Goodrick, 2014). A multiple case design includes all the advantages of a single case study, but extends its advantage in that through repetition of the same analytic procedures on multiple cases, replication can take place, enhancing the validity and consequent potential generalizability of findings (Campbell & Ahrens, 1998, Yin, 2003). Multiple case analyses are useful for the evaluation of intervention programs particularly when understanding the context is important for making sense of the success or failure of an intervention. They are also useful for learning how to tailor interventions to increase the likelihood of achieving the intended outcomes (Goodrick, 2014). In the multiple case study analytic approach, outcomes across many cases can be examined. Researchers can then identify how individual cases are impacted by differing environments or contextual factors, and begin to define contextual conditions under which particular findings (e.g., gains in vocabulary knowledge) might occur.

To explore contextual factors potentially influencing differential response to the VLP reading intervention, with factors organized according to prior defined contextual areas based on Bronfenbrenner's (2005) ecological model, multiple case design was employed. Furthermore,

analyses of the qualitative interview data collected in this research phase, embedded within a broader mixed methods program evaluation study, were focused on multiple-case thematic development using thematic analysis (Braun & Clarke, 2006), to be described shortly. The results of the first research phase (Study 1) revealed differential program effects for some groups of youth. To understand why and how the intervention program was effective for a subset of youth yet ineffective for others, it was important for the second research phase to compare and contrast both groups of youth: Improvers and Non-Improvers. Participants categorized as Non-*Improvers* (i.e., youth who completed the reading intervention but whose outcomes on a key outcome measure did not significantly change following the intervention), represented one cluster of cases. Those categorized as *Improvers* (i.e., youth who completed the intervention but whose reading outcomes significantly improved following intervention) represented a second cluster. To ensure comparability of cases across clusters, cases were matched such that crosscase comparisons accounted for participants' non-verbal ability, age and grade level. The use of a matched comparative sampling approach was essential for elucidating thematic patterns indicating sources of contextual similarity or difference that may be associated with differential response to intervention (i.e., factors that differentiate program responders and non-responders). Descriptive and thematic analysis occurred first within individual cases, then across cases within a cluster, and finally across clusters.

3.2.2 Positioning of the Primary Researcher

Researchers engaged in qualitative analyses often approach their research with their own set of assumptions or beliefs that ultimately influence aspects of their inquiry to some extent (Creswell, 1998). Prior to describing the procedures and analyses employed in this study, it is important to clarify my identity as a researcher, my intentions for the research and personal

investment in the subject area. I am a doctoral clinical psychology student who co-facilitated the VLP intervention across multiple cohorts of its implementation, co-trained tutors who delivered the intervention, and co-developed the intervention program manual – all under the supervision of Dr. Esther Geva. It was important for me to participate directly in this intervention as an occasional on-site facilitator as it allowed for direct interaction with the youth. This in turn allowed for the development of rapport between myself and the youth. I believe this rapport supported their comfort with completing the interviews in this study, particularly given the vulnerability of this population of youth and the potentially sensitive subject matter addressed in interviews. Conversely, I cannot claim that my perspectives of all participants in this study are entirely detached. However, I believe that my personalized experiences with some of these students simply enriched rather than detracted from my findings and interpretations in this study. This qualitative study was developed to fulfill in part the dissertation requirements of my doctorate in School and Clinical Child Psychology at the University of Toronto. Having received extensive training in quantitative and qualitative methodology and more recently, mixed methodology, this research phase reflected an ideal combination of my training in application to the present research problem. My decision to undertake this research was also grounded in my past and present experiences as an advocate for social justice, particularly with respect to culturally responsive education. Thus, this project was viewed as an opportunity to contribute to knowledge about how to improve educational outcomes for culturally and linguistically diverse at-risk youth. As a result, I was particularly invested in the results of this study. My awareness of my own intentions and personal investment in this project continuously implored me to acknowledge and minimize the impact of any biases throughout all steps of the research process (e.g., data collection, analysis, interpretation) so that the validity of findings were heightened. This ascribed value to the preservation of validity in this qualitative phase is consistent with a

positivist approach to case study analysis, as illustrated by the scholars Yin (2003) and Miles and Huberman (1994). Although my training and education has been largely positivist in orientation, and aspects of my study design and analyses are reflective of that, this research is also in part reflective of my increasingly transformative/emancipatory orientation (Mertens 2003, 2007). Contrary to the positivist approach, transformative approaches espouse that knowledge is constantly constructed (rather than reflecting some established truth or reality) for the purpose of improving society, and places significant importance on the lives and experiences of marginalized groups. According to Mertens (2007):

"A researcher can choose quantitative or qualitative or mixed methods, but there should be an interactive link between the researcher and the participants in the definition of the problem, methods should be adjusted to accommodate cultural complexity, power issues should be explicitly addressed, and issues of discrimination and oppression should be recognized" (2007, p. 216).

It is partially a result of this orientation that I chose to conduct a mixed methods study, embracing the opportunity as a primary researcher to give voice to those whose voices are often not sought, or in the worst of times, suppressed. In this way, the hope is that when interventions involving marginalized populations are evaluated and improved upon, the perspectives of those expected to partake in, and hopefully benefit from them are at the very least considered.

3.2.3 Community Context

It is important to consider the community context, including the Pathways to Education program (hereafter P2E, a broader, community-wide initiative) in which participants were simultaneously involved, in order for a comprehensive understanding of the rationale for the methodology, analyses and results derived from this study. A detailed summary of the community context can be found in Study 1.

3.2.4 VLP Intervention

A full description of the VLP intervention is provided in Study 1. The empiricallyinformed, manualized intervention program targeted weaknesses in academic vocabulary and reading comprehension by providing direct and explicit strategy instruction in these areas. The program was implemented as a series of 13 weekly one- and a half hour tutor-led sessions, held after school in a small group setting (e.g., four to six students per group). Each of the 13 sessions was structured identically, beginning with an icebreaker, two reading passages used for strategy instruction, and two games to facilitate consolidation of learning. All program content was delivered by trained tutors -- adult volunteers who received intensive training over the course of two to three weeks in preparation for the delivery of effective, research-based instruction of vocabulary and reading comprehension strategies.

3.2.5 Procedure

This research study was approved by the Research Ethics Boards of Pathways to Education (Regent Park) and the University of Toronto. The present study represents the second phase of a broader, mixed method evaluation of the effectiveness of the VLP reading intervention with marginalized, language-minority and immigrant youth. The initial phase of the study consisted of a pre-post between subjects design involving four sequential cohorts of student participants across four years. The present phase of the study was conceptualized during the time that the second cohort was participating in the intervention. As such, procedures and participant recruitment for this study involved only participants from that cohort. To ensure that

intervention sessions, recruitment for this study was facilitated primarily by a P2E program staff coordinator who was familiar with the students but not directly involved with the study. The P2E staff made initial contact with potential participants, followed by the primary researcher who contacted interested participants to schedule face-to-face interviews. Students were also offered monetary compensation (\$10) for participation in the 45-minute exit interview. After obtaining written consent and assent from students and their parents for participation in the postintervention interview, a total of 29 of 62 participants from cohort 2 met with the author individually in a quiet setting on-site at P2E for the face-to-face interview. Interview respondents included students who were in the intervention and control groups. That is, all participants regardless of their study group (i.e., intervention, control), study completion status (i.e., completed intervention, partially completed) or scores on outcome measures from Study 1 (i.e., improved scores or no change on outcome measures post-intervention) were invited to participate in the post-intervention interviews. The interviewer was blinded to interviewees' scores on outcome measures (until after all interviews were completed), in order to minimize potential biases associated with knowledge of interviewees' measured (vs. reported) reading skill development from pre- to post. Participants were first asked to complete the Pre-Interview Questionnaire (Appendix B), followed immediately by a 30-45 minute semi-structured interview designed to elicit information about their experiences in the program and address other theoretically informed individual and contextual factors that were potentially related (either proximally or distally) to their experiences in the program. All interviews were audio-recorded and transcribed for analysis. Although 29 participants from cohort 2 completed the postintervention interviews that informed the secondary phase of the VLP program evaluation, a smaller purposive sub-sample of these participants were selected for examination in the present

study. Selection was based on their study outcome status as specified by certain pre-determined criteria to be described below.

3.2.6 Measures

Pre-Interview Questionnaire. Demographic information obtained from participants as part of Study 1 (Appendix B) was also utilized in the present study. Participants completed a Pre-Interview Questionnaire developed by the primary researcher in order to provide relevant background information, including: (1) basic demographic data (e.g., age, sex); (2) native language; (3) immigration status; (4) cultural background, and (5) frequency of involvement in extracurricular literacy activities (see Appendix B).

Semi-structured Interview Guide. In order to gain rich insight into student's individual and contextual experiences that potentially impacted their engagement in the reading intervention the primary researcher developed a semi-structured Student Interview Guide. This was done in collaboration with the primary research investigator of Study 1, Dr. Esther Geva (Appendix B). This 30-45 minute interview involved 10-12 open-ended questions that were used as a guide to stimulate discussion of student experiences, with frequent probing questions asked to encourage deep elaboration. Interview questions were grouped according to their contextual relevance in accordance with ecological systems theory (Bronfenbrenner 1997), within five key contextual areas of interest. This approach provided a consistent, theoretical framework for addressing questions across participants (Patton, 2002; Merriam, 1998; Yin, 2003). As such, the interview guide included questions addressing: (1) the individual level context, such as students' academic self-concept, independent involvement in literacy activities, and personal reasons for participating in the intervention program (e.g., How do you feel about your performance in school, like your grades?); (2) the immediate intervention context, such as program attendance

and direct program experiences (e.g., Did you attend every session? What are your thoughts about your tutor?); (3) the peer and (4) family social contexts, for example parental and peer influences on program experiences (e.g., Did your friends encourage you to come to the sessions?), and (5) the broader community and cultural contexts, such as the impact of enrolment in the Pathways to Education community program on VLP experiences and issues around language and cultural identity (see Appendix B for the interview guide). The interviews often but not always addressed the student's perceptions of the relevance of identified broader contextual experiences to their experience in the intervention program (except when direct, interventionspecific questions were asked). All interview questions were exploratory in nature, organized into the contextual categories described here based on evidence from previous research of their potential roles in influencing academic outcomes. The open-ended nature of the interview guide allowed students to flexibly identify and elaborate on those contextual factors most relevant to their educational and life experiences. In addition, the semi-structured format of the interview allowed the researcher to engage flexibly in discourse with participants. Participants were also provided with a list of VLP exercises reviewed during the intervention to facilitate recall of specific activities when the interviewee indicated the need to refer to such information during the interview. All interview recordings were transcribed verbatim by the primary researcher.

Woodcock Reading Mastery Test- Revised. The Passage Comprehension subtest of the Woodcock Reading Mastery Test ([WRMT-R], Woodcock, 1987) was the reading comprehension outcome measure collected in Study 1 that was used to determine which participants would be categorized as 'improved' or 'not improved' post-intervention. A detailed description of the WRMT-R is provided in Study 1.

3.2.7 Participants

Twenty-nine of the 62 students in cohort 2 participated in the post-intervention interviews. Of these, 10 participants (or cases) were purposively selected for the present study (i.e., five *Improvers* and five *Non-Improvers*). It was important to evaluate whether there were any systematic differences at baseline between the 29 post-intervention interview respondents and the remaining 33 non-respondents. Similarly, baseline differences between the subset of 10 selected cases and 19 non-selected cases were examined. The following is a description of the demographic and baseline characteristics of the subset of 29 interview respondents in comparison to the participants who were non-interview respondents (n=33), as well as a description of characteristics of the sample of purposively selected cases (n=10) in comparison to the non-selected cases. This is followed by a description of the process by which the 10 cases that are the focus of this study were purposively selected.

Comparison of interview respondents and non-respondents from Cohort 2. A chisquare test of independence was used to determine whether there were any significant gender differences between the interview respondents (n=29) and remaining cohort 2 non-respondents (n=33). Results indicated that there were no significant gender differences (X^2 =.795, df= 1, p=.372). Independent samples t-tests were conducted to determine whether the two samples differed significantly with respect to age, nonverbal ability, and baseline performance on multiple measures of vocabulary and reading comprehension (descriptions of these measures can be found in Study 1). Results indicated that there were no significant differences between the 29 interview respondents and remaining cohort 2 non-respondents with regard to age [t(60)=0.1.21, p>.05)], nonverbal ability [t(60)=-.24, p>.05)], and baseline (pretest) performance on measures of vocabulary ([EVT-2: t(60)=-1.53, p>.05); PPVT-4: (t(60)=-1.73, p>.05); BWMM: t(60)=-.87, p>.05); BRWI: t(60)=-1.72, p>.05). There was a statistically significant difference on the Gates MacGinitie Reading Test ([GMRT: t(60)=-2.29, p<.05]) reading comprehension measure. Specifically, interview respondents scored significantly higher on this measure at baseline (M=32.92, SD=9.14) compared to the remaining sample (M=25.27, SD=10.17). There was no statistically significant difference between the interview respondents and non-respondents on the other measure of reading comprehension, the WRMT-R [t(60)=-.97, p>.05). Table 16 summarizes the distribution of scores across the two samples on these measures.

Table 16

Characteristics of Interview Respondents and Non-Respondents from Cohort 2

	Interview	Interview non-
	respondents (n=29)	respondents (n=33)
Characteristic (Measure)	M (SD)	M (SD)
Age	14.71 (0.75)	15.24 (1.46)
Nonverbal ability (MAT)	50.67 (9.98)	49.74 (12.07)
Baseline vocabulary		
Expressive vocabulary (EVT-2)	140 (14.78)	127.73 (26.11)
Receptive vocabulary (PPVT-4)	184.84 (16.96)	167.26 (33.50)
Academic vocabulary (BWMM)	17.42 (2.15)	16.27 (4.33)
Vocabulary (BRWI)	16.63 (3.62)	13.92 (4.97)
Baseline reading comprehension		
Passage Comprehension	46.17 (4.99)	43.38 (9.44)
(WRMT-R)		
Reading Comprehension	32.92 (9.14)	25.27 (10.17)
(GMRT)		

Note. Raw scores were used. MAT = Matrix Analogies Test; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; BWMM = Biemiller Words With Multiple Meanings; BRWI= Biemiller Root Word Inventory WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests.*p < .05

Comparison of selected cases and non-selected cases. Fisher's Exact test of

independence was used to determine whether there were any significant gender differences

between the selected cases (n=10) and remaining non-selected cases (n=19). Results indicated that there were no significant gender differences (p>0.05). Independent samples t-tests were conducted to determine whether the two samples differed significantly with respect to age, nonverbal ability, and baseline performance on multiple measures of vocabulary and reading comprehension (descriptions of these measures can be found in Study 1). Results indicated that there were no significant differences between the selected cases and non-selected cases with regard to age [t(27)=-.26, p>.05)], nonverbal ability [t(27)=1.12, p>.05)], and baseline (pretest) performance on measures of vocabulary [(EVT-2: t(27)=0.57], p>.05); PPVT-4: [t(27)=0.26, p>.05)]; BWMM: [t(27)=-.02, p>.05)]; BRWI: [t(27)=-0.27, p>.05]) and reading comprehension ([GMRT: t(27)=-.46 p>.05); WRMT-R: t(27)=0.34, p>.05]). Table 17 summarizes the distribution of scores across selected cases and non-selected cases on these measures.

Table 17

	Selected cases for	Non-selected
	Study 2 (n=10)	cases (n=19)
Characteristic (Measure)	M (SD)	M (SD)
Age	14.8 (0.64)	15.05 (1.16)
Nonverbal ability (MAT)	48.05 (12.61)	51.84 (10.16)
Baseline vocabulary		
Expressive vocabulary (EVT-2)	134.04(19.11)	133 (19.9)
Receptive vocabulary (PPVT-4)	180.02 (21.22)	173.11 (29.33)
Academic vocabulary (BWMM)	16.59 (4.24)	16.57 (3.59)
Vocabulary (BRWI)	14.84 (5.03)	14.46 (4.6)
Baseline reading comprehension		
Passage Comprehension	44.73 (4.85)	44.55 (4.05)
(WRMT-R)		
Reading Comprehension	31.98 (10.24)	28.61 (10.7)

Characteristics of Selected Cases and Non-Selected Cases

(GMRT)

Note. Raw scores were used. MAT = Matrix Analogies Test; EVT-2 = Expressive Vocabulary Test, 2^{nd} Edition; PPVT-4 = Peabody Picture Vocabulary Test, 4^{th} Edition; BWMM = Biemiller Words With Multiple Meanings; BRWI= Biemiller Root Word Inventory WRMT-R = Woodcock Reading Mastery Tests – Revised; GMRT = Gates-MacGinitie Reading Tests. *p<.05

Additionally, it is important to note that analyses conducted in Study 1 indicated that this overall cohort from which participants were interviewed (i.e., cohort 2), did not differ significantly from the other three cohorts of participants in the larger study (refer to Table 5 in Study 1 for a summary).

3.2.7.1 Procedure for purposive selection of cases

Of the 29 interview respondents, a total of 10 were purposively selected to form individual cases in this study. Each of the selected participants was defined as an individual case, representing the unit of analysis (Yin, 2003). To better understand why and how the intervention was effective for some participants and ineffective for others, it was important to examine two independent case clusters: (a) *Improvers*, for whom the intervention was effective with respect to at least one primary reading outcome evaluated in the first study, and (b) *Non-Improvers*, for whom the intervention was ineffective with respect to positively changing a specific reading outcome of interest from the first study. The rationale for the use of two comparative case clusters was to support the exploration of sources of contextual similarity or difference that may be associated with differential responses to the reading intervention.

The participants or cases in this qualitative follow-up study were thus carefully selected based on the quantitative findings in Study 1. Cases that formed the *Improvers* cluster were selected first, followed by cases that formed the *Non-Improvers* cluster. Specifically, the application of five sequential selection criteria allowed for the careful selection of cases for each cluster. In order to represent an eligible case for this study, it was required that each participant: 1) was part of the intervention group; 2) completed at least 9 of 13 intervention sessions (and thus categorized as a program completer); 3a) demonstrated statistically and clinically reliable improvement (Jacobson & Traux, 1991) measured by pre- to post change scores on the WRMT-R reading comprehension task, a primary intervention outcome (for inclusion in the *Improvers* case cluster) or 3b) demonstrated no statistically or clinically significant change on the WRMT-R (for inclusion in the *Non-Improvers* case cluster); 4a) had a standardized residual gain score on WRMT-R that was ≥ 1 SD above the group mean (for retention in *Improvers* cluster) or 4b) had a standardized residual gain score on the WRMT-R that was on the WRMT-R that was within ± 1 SD of the group mean, and 5) met additional inclusion criteria based on age, grade, nonverbal ability and the specific after-school tutoring group they were placed in, to serve as a paired match for between-group case comparative analyses. These case selection criteria as applied to this study will be further described below.

Criterion 1. The first step in determining case eligibility was to exclude control participants from the total of 29 interviewees. As a result, three cases were deemed control participants as defined in Study 1 (i.e., participated in less than two intervention sessions or was assigned to the control group upon study enrolment). These cases were thus ineligible for selection (see Figure 4 for a breakdown of resultant sample sizes per criterion applied).

Criterion 2. As defined in Study 1, intervention completers were defined as participants who completed at least 9 or more of the 13 intervention sessions. Of the 26 remaining cases, four cases were categorized as non-completers and subsequently ineligible, leaving 22 cases for further consideration.

Criterion 3: Reliable change. Given the results of Study 1 indicating that the intervention was not uniformly effective in improving participants' vocabulary and reading comprehension

outcomes, it was important to utilize an index of change or remediation that considered normative population scores, as it can serve as an alternate powerful indicator of postintervention change (Jacobson & Traux, 1991). This is consistent with Fuchs' (2003) recommendation for use of a normalized post-treatment status as a cut-point for distinguishing intervention responders from non-responders. Few studies report remediation rate of participants following intervention. The remediation rate is an index of the proportion of participants whose post-intervention scores were both statistically and clinically improved. It is assessed using a standardized methodology that takes into account fluctuations in test scores over time (i.e., testretest reliability, regression to the mean) (Jacobson & Traux, 1991) and is thus less susceptible to these potential measurement errors. This is particularly important given the limitations of the findings of Study 1, where learning effects and other test-retest issues potentially minimized the likelihood of detecting intervention effects. Accounting for such problems allows for increased confidence in detecting cases of valid improvement among participants. Clinical significance of change was computed following the Jacobson and Truax (1991) criteria for the outcome measures at post-test (using standard scores) in order to determine remediation rates of participants. According to Jacobson and Truax (1991), a participant can be considered remediated if criteria for both *statistically reliable* change and *clinically significant* change are met. Statistically reliable change occurs when the change found between pre to post reflects actual change that is not a result of measurement error of the instruments used. Clinically significant change occurs when participants meet clinical cut-offs on norm-based, standardized measures at post-test such that their post-test score is statistically more likely to be drawn from the nonclinical rather than the clinical population (Wise, 2004). In order to calculate remediation rates, test-retest reliability and clinical cut-off data of the measure is required. This information was only available for the standardized outcome measures used in Study 1, namely two

vocabulary outcome measures (Expressive Vocabulary Test [EVT-2, 2007]; Peabody Picture Vocabulary Test [PPVT-4, 2007]) and one measure of reading comprehension (Woodcock Reading Mastery Test-Revised, Passage Comprehension subtest [WRMT-R, 1987]). Of these three, the WRMT-R measure was chosen as the sole outcome measure for case selection to address the current study goal of assessing factors differentiating reading comprehension development, the ultimate goal of reading. Additionally, it was beneficial to utilize a single outcome measure to determine case cluster categorization as it heightened the reliability of crosscase comparative analyses. The use of multiple measures for joint consideration in the case selection process was also avoided due to the potential for significantly minimized sample size. For the WRMT-R, the clinical cut-off was a standardized score at post-test less than 90 (Fuchs, 2003). Scores below this point fall within the Below Average range relative to a normal population (Woodcock, 1987). Thus, standardized scores were required to have increased at posttest to a SS value of 90 or higher to be deemed clinically significant change (improved scores). Scores that were below this cut-off were deemed clinically unchanged. Statistically reliable change was determined by calculating the reliable change index (RCI) from pre to post (Jacobson & Truax, 1991). To calculate the RCI, the post-test scores were subtracted from the pretest scores and then divided by standard error of difference between the two test scores (Sdiff). The standard error of difference was determined from the standard error of measurement, SE: $Sdiff=\sqrt{2}(SE)2$. SE=s1 $\sqrt{1-rxx}$, where s1 represented standard deviation of the WRMT-R measure at pre, and rxx represented the test-retest reliability of the measure. RCI scores of 1.96 or higher were considered statistically reliably improved scores (p < .05). After jointly applying the criterion of statistically reliable and clinically significant change to the 22 cases, five cases met criteria for categorization in the *Improvers* cluster, while 15 cases met criteria for the *Non*- *Improvers* cluster. Two cases met criteria for *reliable deterioration* (i.e., their change scores indicated significant change in the negative direction) and were deemed ineligible.

Criterion 4. The previous criterion required cases to be measured up against an external, normative standard of change. Criterion 4 required cases to be compared to the sample mean to verify that relative change scores of retained cases remained within specified parameters to facilitate between group/cluster comparisons. Specifically, to further identify cases based on how they compared relative to the mean change scores of the entire sample of intervention completers interviewed in this study phase (n=22), cases with standardized residualized gain scores on the passage comprehension task of the WRMT-R that were ≥ 1 *SD* above the group mean were retained as *Improvers*, while those with gain scores within ± 1 *SD* of the group mean were retained as *Non-Improvers* (see Figure 4). As a result, the five cases previously identified as *Improvers* remained, and only 14 of the 15 *Non-Improvers* were retained.

Criterion 5. Yin (2003) recommends a range of six to ten cases for multiple case designs. The primary researcher decided to retain a total of 5 cases per group, for a total of 10 cases. It was also decided that each case in the *Improvers* category would be matched to each case in the *Non-Improvers* category by age, grade and nonverbal ability (to better control for factors potentially differentiating the groups). Therefore, each of the remaining 14 *Non-Improvers* was subjected to this further selection criterion of case matching. Six cases could not be matched due to incompatible age, grade and/or nonverbal ability. Of the remaining 8 cases (each being a viable match for at least one of the five cases in the *Improvers* group), five were selected based on their proximity in age (in months) to the matched counterpart in the *Improvers* (n=5) and (2) *Non-Improvers* (n=5). Table 18 provides a summary of the demographic characteristics of the two groups of cases. Independent samples t-tests revealed that mean differences between the

groups were non-significant, p>.05. A chi-square test of independence indicated that there were no significant gender differences (X^2 =.33, df =1, p=.57). Each case was assigned a pseudonym to maintain anonymity. The participants attended an average of 11.5 VLP intervention sessions (*SD*=1.51).



Table 18

Characteristics of Improvers and Non-Improvers

	Improvers (n=5)	Non-Improvers (n=5)
Characteristic (Measure)	M (SD)	M (SD)
Age	14.80 (.64)	14.80 (.64)
Nonverbal ability (MAT) Standard		
Score	90.38 (5.28)	88.00 (4.00)
Baseline vocabulary		
Expressive vocabulary (EVT-2)	129.70 (18.80)	138.38 (19.41)
Receptive vocabulary (PPVT-4)	177.40 (16.24)	182.63 (26.12)
Academic vocabulary (BWMM)	16.60 (2.19)	17.38 (2.62)
Vocabulary (BRWI)	15.70 (3.33)	17.00 (3.73)
Baseline reading comprehension		
Passage Comprehension	42.20 (3.63)	47.25 (6.07)
(WRMT-R)		
Reading Comprehension	32.20 (6.42)	31.75 (11.39)
(GMRT)		

3.2.8 Data Analysis

Data analysis within a case study design is a process of making meaning out of the data (Merriam, 1998; Yin, 2003; Stake, 2005). This process involves "consolidating, reducing and interpreting what people have said and what the researcher has seen and read" (Merriam, 1998, p.178). Multiple case analysis specifically involves a combination of within case and cross-case

analytic methods (Yin, 2003; Creswell, 1998). Overall data analyses in this study involved an application of Stake's (2005) multiple case analytic approach of "categorical aggregation" in combination with thematic analysis as described by Braun and Clarke (2006) to ultimately identify overarching cross-case themes. The following sections provide a brief description of each of these approaches followed by a description of the specific analytic steps undertaken in this study.

Theoretical Thematic Analysis. Theoretical thematic analysis (Braun & Clarke, 2006) involves searching for repeated patterns of meaning across data. The analysis, guided by research questions, is driven by a theoretical framework. This allows for the generation of themes around a pre-specified topic area (e.g., an ecological context factor in this study) such that the derived themes expand on the broader, pre-defined framework. In this study, themes were generated at the semantic level, such that the analytic process progressed from description (i.e., organization of data to display patterns in semantic content) to interpretation (i.e., examination of the significance of the patterns, their meanings and interpretation) in relation to existing literature (Braun & Clarke 2006). Braun & Clarke (2006) delineate six steps in completing thematic analyses including a) familiarizing oneself with the data; b) initial coding; c) searching for themes; d) reviewing themes; e) defining themes and f) producing a report. Thematic analysis occurred throughout all stages of analysis in this study.

Categorical Aggregation. Although thematic analysis was the underlying analytic approach used to reduce the data into broader themes, this study also employed Stake's (2005) categorical aggregation approach to more directly facilitate cross-case analyses. Categorical aggregation involves clustering the data into categories by searching for 'multiple instances' in the data from which new meanings or cross-case assertions can emerge. This is accomplished by dialectical review and comparison of cases (e.g., reviewing a case while simultaneously holding

the other cases in mind to help make sense of emerging differences or similarities) as well as identification of atypical cases or counterevidence. Yin (2003) emphasizes the utility of this approach of identifying similarities and differences in maximizing the validity of findings.

The following specific analytic steps were taken in this study in a recursive rather than linear manner, where the primary researcher moved flexibly and repetitively back and forth across the different stages as needed to make sense of the data (Braun & Clarke, 2006). It is important to note that the various stages of analyses complemented one another. Cross-case analyses contributed to the identification of themes and thematic relationships that characterized a broad range of participants' experiences. Simultaneously, within case analytic steps made it possible to develop themes in a way that carefully accounted for the specific factors that shaped individual participants' experiences. Table 19 summarizes the various within- and across-case analytic strategies used.

Table 19

Within- and Across-Case Analytic Strategies

Comparison	Purpose	Strategy	Product
Within individual participants	Identify important aspects of individual students' experiences	Close reading of individual interviews and narrative summaries	Coding categories, preliminary themes
Across individual participants within a case cluster (e.g., all <i>Improvers</i>)	Identify thematic variation and configuration of themes within all <i>Improvers</i> and <i>Non-</i> <i>Improvers</i> separately;	Data coding and display	Coding categories, subthemes
Across clusters (i.e., Improvers vs. Non- Improvers)	Compare thematic variation across <i>Improvers</i> and <i>Non-Improvers</i> ; identify cross-case contextual factors	Data coding and display; review of within and across- case summaries; comparison of matched dyads	Cross-case subthemes; superordinate (meta) themes

<u>Stage 1:</u>

Following verbatim transcription of audio recordings of interviews, data from the preinterview questionnaire was integrated with interview data. The transcribed interviews were then loaded onto Nvivo 9.0, an electronic qualitative data analysis software program. I became familiarized with the data by reading all the interviews, actively noting preliminary thoughts and patterns as I read through them. I re-read the interviews of all ten cases while assigning initial codes to the text (Braun & Clark, 2006) using Nvivo. Codes were broadly organized around the six primary contextual areas theorized to be implicated during reading comprehension skill development (i.e., individual, intervention, peer, family, community and cultural contexts). Codes represented the noteworthy comments of the youth that held the potential of later serving as the basis for repeated patterns or themes. Coding occurred systematically, in an iterative, comparative manner across the data set. An initial list of 47 codes was generated at this stage. Further thematic analysis was conducted on the generated codes, where codes indicating similar ideas were aggregated or sorted into a broader category within each contextual area. A list of 32 coding categories across the six contextual areas was generated.

Stage 2:

Further analysis in this stage involved visually displaying the coding categories and the data supporting them from each individual case in a manner that allowed for initial thematic analyses within cases, across cases in the same group (e.g., across all *Improvers*) and between the two groups of cases. Specifically, coded text segments relevant to each coding category from Stage 1 were charted per participant (and per contextual area of interest) in a series of word tables (Yin, 2003). Word tables display data from individual cases, organized according to a common theoretical framework (Yin, 2003). This allowed for visual comparison of coded

segments associated with all cases of *Improvers* and *Non-Improvers* at once, organized by ecological context of relevance (e.g., individual context, community context). As recommended by Miles & Huberman (1994), this charting process allowed for thematic analysis of the overall patterns in the word tables, resulting in the generation of the first set of preliminary within case, within group and between group subthemes. These subthemes were integrated with the subsequent analytic phases to contribute to cross-case conclusions about the differential response to the intervention of *Improvers* and *Non-Improvers*.

<u>Stage 3:</u>

The remaining stage involved a series of more in depth thematic analysis steps, integrating Stake's (2005) recommended process for multiple case study analysis in efforts to develop subthemes first within cases, then within each group, and ultimately between groups. The primary research question remained central: identifying the contextual factors associated with differential response to intervention among the two carefully selected groups of cases. With this ultimate goal in mind, along with a theoretical framework guiding organization of themes, the following steps were taken.

1) <u>Within case analysis:</u> It was important to understand individual cases in depth prior to cross case analysis (Stake, 2005). Transcripts from each interview were reviewed and a synopsis or narrative summary of each case was developed by reading each interview and making notes of relevant findings or potentially thematic areas of interest. A thematic area from each new interview that was sequentially read was constantly compared to other cases for confirming or disconfirming evidence. Each case was constantly reconsidered in light of learnings from other cases - that is, each was read while the researcher held findings from the other cases simultaneously in mind. Case narratives were written up in a structured format that incorporated both interview and questionnaire data gathered from this phase of the study. The researcher

highlighted aspects of each case thought to potentially inform cross case analysis. The purpose of developing these individual case narratives was to form an evidence base for subsequent group-level analysis; as such, case narratives were not reported as independent findings.

2) <u>Within-group cross-case analysis:</u> The case narratives, essentially the themes of the cases in the *Improvers* group, were compared with one another in an iterative, reflexive manner, noting similarities in themes as well as cases that differed in any way (Yin, 2003). This process was repeated with *Non-Improvers*.

3) <u>Between group cross-case analysis:</u> An important step of between group analyses involved comparing the content and making note of important themes emerging across matched dyads between the two groups. Five dyads were compared, where each *Non-Improver* was matched to an *Improver* of similar age, grade, and nonverbal ability. Each dyad was analysed separately, via ongoing thematic analysis to contribute to and refine previously identified cross-case themes. Similar and differentiating contextual factors within dyads were noted. As with previous analytic steps, and in concordance with Stake's (2005) method, a list of between-group themes (according to systematic dyadic comparisons) and their descriptions was generated.

4) <u>Further refinement of cross-case analyses</u> involved inductive analysis using themes established in the previous step. Specifically, this step entailed the identification of similar and divergent themes iteratively across all five matched dyads until between group cross-case themes were generated. Any unique learning offered by specific dyads that were not replicated in other dyads was especially noted. New themes that were generated at this stage elucidated other areas that potentially differentiated *Improvers* from *Non-Improvers* when age and grade level were accounted for, helping to rule out the influence of these characteristics on the exploration of factors that may be associated with response to intervention. 5) Integration of themes: The thematic integration process involves expanding, merging and creating thematic categories. Thematic synthesis was initiated whenever emergent themes across cases suggested broader concepts to be considered (Miles & Huberman, 1994). As I progressed through the iterative coding process, what emerged were areas where participants agreed and disagreed, recurring themes, and categories of themes. Thematic synthesis occurred in an ongoing manner at each of the analytic stages given the reflexive and iterative nature of the analytic approach taken (Stake, 2005). This process of thematic integration was intensified at the last stage, as the broadest thematic categories representing the data were developed through careful consideration of thematic results from prior analytic steps. Ultimately, themes that were common across cases were retained rather than individual case-based themes. Thus, based on the readings, notes and thematic categories generated at previous steps, a table was created summarizing 11 key broad thematic findings in this study. These final themes represented the factors that were unique and similar among *Non- Improvers* and *Improvers* for each of the six contextual factors and that potentially were associated with differential response to intervention.

3.2.9 Assessing Quality of Findings

There are numerous guidelines available for the assessment of the quality and validity of qualitative research (Creswell & Plano Clark, 2011; Kitto, Chesters, & Grbich, 2008; Yin, 2003; Lincoln & Guba, 1985). Adhering to such guidelines can help strengthen the rigor of the study. Creswell & Plano Clark (2011) outline the following methodology-based criteria for assessing quality of a qualitative study: 1) rigorous data collection; 2) philosophical assumptions framing the research are outlined; 3) use of an accepted approach to inquiry (e.g., case study); 4) use of validity strategies to confirm accuracy, and 5) multiple levels of data analysis conducted. Table

20 highlights the methods employed in this study as recommended by Creswell and Plano Clark (2011), to ensure quality of the present study.

Table 20

Assessment of Research Quality

Creswell & Plano Clark (2011)	Strategies Employed to Meet Criterion
Criterion	
Rigorous data collection	Sampling, data collection and analysis
	-use of small, purposive sample
	-use of open-ended semi structured questions
	allowed for flexibility in responses along a consistent
	framework
	-collected multiple forms of data (open-ended
	interview, questionnaires)
	-data was prepared for analysis using a qualitative
	data analysis software program
Philosophical Assumptions Framing the	-Primary researcher's philosophical leanings and
Research Outlined	related assumptions that framed data collection and
	analysis were explained and their impact considered
Use of an acceptable approach to inquiry	-Identified and defined analytic approach of multiple
	case analysis
	-Procedures for conducting multiple case analysis (in
	a manner consistent with existing literature) is
	clearly presented, with features of three prominent
	approaches (Braun & Clark, 2006; Stakes, 2005;
	Yin, 2003) applied
Use of validity strategies to confirm	-Used multiple procedures for validation
accuracy	-Used triangulation (of multiple methods and data
	sources) to corroborate data and findings
	-Data analysis process was iterative, allowing for
----------------------------------	--
	progressive validation
	-Divergent findings identified and incorporated to
	ensure that rival explanations have been considered
	(Yin, 2003)
	-Internal Validity: use of a case-matching process
	helped to control for potentially confounding
	variables
	Reliability: Details about study content and process
	were provided, to allow for potential replication
Multiple levels of data analysis	-Multi-level analyses were exemplified through the
	progression from specific codes or themes to broader
	themes, to even more abstract key themes

3.3 Findings

There were no separate sections in this study devoted to findings of individual cases as the focus was the integration of cross-case findings. Individual cases served as the evidence base for cross-case analysis from which cross-case themes were organized around pre-determined contextual areas as indicated below. Within each theme, appropriate and representative examples from the various cases were drawn (Yin 2003). Each case was assigned a pseudonym to maintain anonymity.Wherever applicable other identifying information was also replaced with pseudonyms (e.g., School X, Tutor Y).

As discussed earlier, the overarching conceptual framework of thematic organization in this study was based on Bronfenbrenner's (2005) ecological model. Thus, final themes were developed for each of four contextual areas: individual context (3 themes), VLP context (intervention context; 3 themes), social context (i.e., peer and family contexts; 2 themes), and broader community context (i.e., school, community and cultural contexts; 3 themes). Themes at the individual contextual area were: 1) achievement orientation; 2) academic self-concept, and 3) sense of future aspirations. Themes at the intervention level were: 4) personal investment in positive outcomes; 5) resourcefulness and 6) tutoring group climate. Within the social context (peer and family contexts), themes were: 7) reliance on peers and 8) nature of parental support. Within the broader school, community and cultural contexts, themes were: 9) perceptions of school; 10) engagement with school and broader community and 11) sense of cultural identity. In summary, the within and across case analyses yielded 11 themes across four contextual areas reflecting dimensions of factors potentially differentiating the post-intervention outcomes of youth.

As a result of an intensive thematic integration analytic process, one super-ordinate theme emerged, representing the ways in which the themes in each contextual area were connected with one another to form a cohesive narrative of the factors that potentially differentiate youth whose reading outcomes improve over time and those whose outcomes remain unchanged. This theme of *personal agency* spanned across the four major contextual areas to relate the predominant themes. All thematic findings are summarized in Table 21.

Table 21

Contextual Area	Theme	Description
Individual Level	Achievement orientation	Extent to which participant was driven to excel academically
	Academic self-concept	Participant's perception of their overall academic performance
	Sense of future aspirations	Sense of future career or educational goals
Intervention Level	Degree of personal investment in positive intervention outcomes	Extent to which participant was behaviorally and affectively invested in using the program as a tool to improve skills

Summary of Themes

	Tutoring group climate	The overall behavioral and affective tone of the tutoring group
	Participant resourcefulness	Extent to which participants independently sought out opportunities to maximize their learning during sessions
Social Level (Peer & Family)	Reliance on peers	Extent of peer reliance for supporting participant engagement with the VLP or general peer support of academic work
	Nature of parental support	The quality of parental (and familial) motivation and encouragement as well as the extent of parental involvement in participants' academic activities
School, Community & Cultural Level	Perceptions of school	The nature of participant's views of their school environment and experience
	Engagement with school and broader community	Extent of participant engagement with school community and broader community via extracurricular activities and volunteer work respectively
	Sense of cultural identity	Participants' sense of personal identification with a particular ethno-cultural identity or identities
	Integrative Theme: Personal Agency	The participants' capacity and propensity to take purposeful action in exerting influence over their learning environment

Individual-level context

Theme 1: Achievement orientation

Achievement orientation as a theme referred to the extent to which a participant was achievement-oriented or driven to excel academically. Although both Improvers and Non-Improvers generally spoke about being motivated to do well in school, Improvers tended to exhibit a higher level of achievement orientation. Improvers tended to spontaneously comment on their high grades, and were objectively high-achieving based on their reported grades in core subject areas. For example, in discussing his academic performance, Wajid expressed that his grades are so high that he is "on the high honor roll" which is a "higher honor…like 90% average." Anwar, another Improver, also shared that he does not have any difficulty obtaining high marks in Science, citing a grade of "95% plus" and that he is "at the top of the grade in the school."

Some Non-Improvers also commented on their academic grades, reporting average level performance relative to their peers. None reported below average grades. Javitha, a Non-Improver, shared that she feels that her grades are "average." Naya, another Non-Improver, shared: "My average is like over 70s. I feel like I can improve." One Non-Improver, Sarah, described herself to be on the honor roll unlike the other Non-Improvers. However, the manner in which she described her grades demonstrated poor confidence in her skills: "I mean like I'm doing really bad in Math, and for Grade 11 I really want to bring up my marks... I'm in the honor roll with an 83 so it's not that bad..." While Non-Improvers did not make any mention of receiving academic awards, Improvers (60%) shared that they received academic awards across a span of subjects such English, Math and Science. Improvers were also more likely to comment about their high academic placement relative to their same-grade peers. Another exemplar of Improvers' high achievement orientation was the fact that they were more likely to take initiative and be independent in seeking out tutoring or other academic support when needed. Non-Improvers were more likely to make statements indicating a tendency to settle for available resources (e.g., Pathways to Education tutoring, internet, friends). Improvers also tended to be those participants who placed a priority on their academic work despite their involvement in extracurricular activities, whereas Non-Improvers reported that their academic work was sometimes compromised in the pursuit of extracurricular activities.

Theme 2: Academic self-concept

Related to achievement-orientation was also the theme of academic self-concept, defined in this study as participants' perception of their overall academic performance. Improvers were more likely to make statements reflective of a positive academic self-concept, in comparison to Non-Improvers who made less positive statements. Grace, a grade nine participant, and Anwar (Grade 11) expressed confidence in their skills to the extent that they often helped their peers with academic work:

In general I help my friends when they need it, like if we're in the same class I stay back with them, study, encourage them and help them finish their work cause like I take my work seriously...I'm actually doing better than I thought I would be, I'm actually happy with my grades where they are now...so yeah I'm feeling good." (Grace, Grade 9)

I like tutor for other subjects at my school, and a lot of my friends come in for help...Well compared to other people I feel I'm doing really good. (Anwar, Grade 11)

On the whole, Non-Improvers exhibited a more neutral self-concept, reflected by inconsistent and sometimes negative statements about their academic performance. For example, Naya explained: "I feel like I could do better, I'm happy, I'm kinda sad. I feel like I should...like if I tried harder I could have like got to a higher point" (Naya, Grade 11).

Theme 3: Sense of future aspirations

Participants' future goals and aspirations were an important differentiating factor between Improvers and Non-Improvers. Although participants in both groups described an emerging sense that they wished to complete postsecondary education and become involved in a career, Non-Improvers did not present clear or well-articulated ideas regarding their future aspirations. Improvers, however, were more likely to have firm ideas about their future career goals, which often entailed professional careers. Improvers, for example, stated a desire to "get into medical school" (Anwar, Grade 11), "go to University" (Wajid, Grade 9) or to become "an immigration officer or legislator" (Amala, Grade 9). Two of the Non-Improvers expressed interest in the careers of a "nurse" and "social worker," however they both indicated uncertainty about these possibilities. Two other Non-Improvers did not articulate any future goals. One made the general statement that he wished to work "somewhere in the business field," yet did not elaborate on a specific career of interest.

VLP (Intervention Level) Context

Themes at the intervention program level were the degree of personal investment in positive intervention outcomes, the tutoring group climate and resourcefulness of the participant.

Theme 4: Personal investment in positive intervention outcomes

Compared with Non-Improvers, Improvers were more invested in their own selfimprovement, manifested through their initial motivation to participate in the program, their rationale for persisting with the program and their engagement with the program content.

Initial motivation to participate in the program. Improvers' cited reasons for participating in the program exemplified their investment at the outset in improving their reading skills. When discussing their motivation to participate in the intervention, Improvers shared that they perceived the program as a means to achieve their personal academic goals either in the subject of English or reading in general, and expressed a desire for ongoing self-improvement. Improvers also expressed that they were motivated by incentives that had the potential for a lasting impact, involving less immediate gratification (e.g., obtaining volunteer hours to meet high school graduation requirements, meaningful experience, opportunity to add their participated "for volunteer hours, experience, and it would look good on my resume, and just to improve more on my reading" (Wajid, Grade 9). Similarly, another Improver commented:

I think it was a good program because they taught you about vocabulary and I think I kind of needed that. My mom wanted me to go because she was like you should get your vocabulary straight, I don't know I thought it was beneficial to go there and you get volunteer hours too so that was a good program. (Amala, Grade 9)

Non-Improvers however, did not indicate that they were particularly invested or interested in their learning outcomes post-intervention based on their expressed reasons for program participation. Non-Improvers tended to either describe the program as a service that was developed solely for the gain of others (e.g., for the research team) or did not articulate an intended program purpose. Participants in this group also were also more likely to describe being in part motivated to participate in the intervention based on factors that were distal to their own self-improvement or to key program goals -- even in cases where the desire to improve English skills was also cited as a motivator. For example, these participants described being motivated by the possibility of obtaining snacks, token incentives (e.g., rewards/prizes), financial incentives or opportunities for social interaction. Grade 9 participants Javitha and Priya both expressed that they participated because it would be "fun." Sandra, a Grade 9 participant, recalled that her initial interest in participating was based on persuasion from Pathways to Education staff as well as the potential for remuneration:

"...so everybody's like "oh money, ok let's do it!" Like we wouldn't have done it if...she kinda persuaded us she's like "come on!" and once we got into it we were like oh well we're already here, you know? (Sandra, Grade 9).

Rationale for persisting in the program. Another way in which both groups of participants indicated their personal investment in positive learning outcomes was through their expressed rationale for remaining in the program for its duration. Improvers were more likely to describe that they remained in the program for the purposes of contributing to their own learning, despite occasions where they felt inclined to be absent. For example, Nadia shared that she continued to come because she was "learning new things," even though "some days I'd come and I wouldn't be in the mood for it but I'd still come" (Nadia, Grade 9). One Improver expressed feeling uncertain that the program would meet his learning needs, explaining that instead of leaving the program, he found a way to ensure that his learning needs were met:

The point of the program was to teach the students new vocabulary and stuff, so I used words from my homework and I learned. Like at first I felt that it wasn't enough for me, so I tried to solve that problem by not leaving it. So I found my solution. (Anwar, Grade 11).

Non-Improvers demonstrated much less personal investment in positive learning outcomes based on their cited reasons for persisting with the intervention program. Specifically, when discussing their rationale for persisting, Non-Improvers focused on factors such as personal entertainment, opportunities for socializing with peers, possibility of receiving prizes, having friends in the program, being encouraged by their tutor to remain in the program, and resignation to persist based on the fact that the participant was already enrolled. For example, when the researcher explored her reasons for remaining in the program for its duration, Sandra noted: "I didn't want to be there, but then I already like did it, so I'm like I might as well." Another Non-Improver commented that her best friend was "the one who persuaded me into staying... but if she wasn't there I would definitely not have come" (Javitha). It is important to note however that despite their reasons for remaining in the program, 80% of the participants across both groups indicated that they made concerted efforts not to miss sessions unless they needed to complete urgent schoolwork such as preparation for exams or project completion.

Engagement with VLP content. Improvers were more actively engaged with the program content and learning components. This engagement with the content (and consequent personal investment in positive learning outcomes) was reflected in the differential manner in which Improvers and Non-Improvers described program activities and their favored (or less favored) program aspects. When describing program content, Improvers were more likely to reference specific components and strategies that they found beneficial or non-beneficial to their learning. They were also more likely to identify program strategies that they learned and planned to continue to apply post-intervention. With the exception of one individual, all Improvers named

specific newly acquired strategies that they were currently applying when reading. When discussing their favored program aspects, Improvers focused on highlighting the utility of program materials (e.g., selected readings, worksheets), their dedication to learning new strategies and the personal benefits of improving their reading skills. Improvers were also more likely to make suggestions for improvement of specific content areas. Improvers' program suggestions primarily involved providing instruction tailored to student ability levels and providing a broader variety of texts from which students could choose for the purpose of reading instruction. They suggested that these texts be specifically tailored to individual students' reading interests. In describing their tutors, Improvers' narratives overwhelmingly focused on tutor behaviors that were directly related to delivery of program content (e.g., tutor's application of specific strategies, ability to facilitate participant understanding of instruction), indicating their attentiveness to key program components that were important for driving improved outcomes. Improvers also were more likely to explicitly express that they retained some of the content they had learned as a result of the intervention, describing certain aspects of their learning as forms of stable and enduring knowledge.

In contrast, Non-Improvers demonstrated that they were generally less engaged with the content and thus minimally invested in positive program outcomes. When describing program content, Non-Improvers' comments were more often focused on program aspects distally related to key outcomes, such as a focus on playing games, making friends, enjoying snacks and having "fun." Non-improvers' narratives around favored program aspects were often limited to these perceived aspects of enjoyment rather than on the experience of learning vocabulary and reading comprehension strategies. In fact, Non-Improvers sparsely mentioned specific strategies. Similar to Improvers, some Non-Improvers made suggestions for improving the program that included requests for a broader variety of text materials. However, the majority of suggestions made by

Non-Improvers reflected their lowered attentiveness and engagement with core program content, with a focus instead on program structure issues (e.g., session length, program duration, group size). Some Non-Improvers noted that the program content was "boring", repetitive and not sufficiently difficult, as did some Improvers. However unlike the Improvers, the Non-Improvers did not express any direct personal efforts made towards remedying the fact that they were not getting the most of the learning opportunity. Rather, these participants continued to engage with the program despite sentiments that their learning was not challenged, and in some cases, that any new learning was not sufficiently retained. The following excerpt from an interview with Sandra, a Non-Improver, highlights this latter finding:

Interviewer: Are there any lessons you've learned that you are actually using from the VLP?
Sandra: Well I dunno, maybe not consciously, maybe I just do it. I don't know...I can't remember a lot of things we did, except the games and stories. (Sandra, Grade 9)

Additionally, unlike Improvers, Non-Improvers tended to provide vague, generalized descriptions of tutor characteristics that were not as closely related to the instruction they received (e.g., tutor's friendliness). Comparison of matched pairs of participants from each study group, particularly those who were in the same tutoring group, solidified the pervasive nature of these differences in engagement with program content among Improvers and Non-Improvers. Wajid (Improver) and Priya (Non-Improver), for example, were in the same tutoring group and yet embodied these described differences in their extent of engagement with program outcomes and consequently the extent of personal investment in positive outcomes.

Theme 5: Tutoring group climate

The tutoring group climate, defined as the overall behavioral and affective tone of the tutoring group, was another factor at the program contextual level that differentiated the group of program Improvers from the Non-Improvers. In total, six different program tutors were

represented across the ten cases that were evaluated in this study. Notably, all but one of the participants (a Non-Improver) described having generally positive experiences of their tutors and tutoring groups. Improvers noted that they benefited from being members of tutoring groups that they perceived to be cohesive, cooperative and coordinated in terms of focusing their group efforts on learning the outlined program strategies. For example, some Improvers described the aspects that they particularly favored in their tutoring groups in the following way:

None of them were like discouraging or anything. They always had full attention and they all participated and they all attended every session. (Wajid, Grade 9)

They were just like, they're not loud or rowdy, we don't speak over each other, we take turns, and stuff like that. So it was nice working with them. (Nadia, Grade 9)

We worked together, we got everything done....everything went smoothly. And every single time we at least learned something. (Anwar, Grade 11)

Non-Improvers described being members of tutoring groups that were often marked by poor cooperation among its members, poor behavioral and affective engagement (e.g., tardiness, poor attendance, negative attitudes towards program activities), and a focus by peers on "having fun" rather than directed efforts to achieve learning goals. Sandra, for example, described that her group's morale was impacted by some members who were not interested in engaging with the activities or even remaining in the program: "They were always complaining, saying "when can we leave?" and "this is so uhh". They weren't appreciative." (Sandra, Grade 9). In some cases, these participants directly attributed the overall tone of the group to their own diminished engagement and learning.

Theme 6: Participant Resourcefulness

One common complaint expressed by both Improvers and Non-Improvers was that many of the program strategies such as the Know-Want-Learn (KWL) chart (see Pasquarella et al., 2013 for a description of program activities) were already familiar to them prior to their

involvement in the program, and did not constitute new learning. As such, some participants reported that they felt their learning was "repetitive" and "boring." In light of this finding, another predominant theme with respect to the VLP program context emerged. This was the theme of participant resourcefulness - the extent to which participants independently sought out opportunities to maximize their learning during sessions. Improvers were more likely to exhibit resourcefulness, in that they actively worked to adapt program content to complement their prior learning experiences and to meet their personalized learning goals. Improvers accomplished this in a variety of ways. Some directly informed their tutors of their learning needs, and worked together with their tutors to identify instructional modifications that would best suit them (e.g., reading more complex texts to identify more advanced vocabulary words; using their assigned school texts as the material for vocabulary learning so that instruction was related to academic work; modifying the application of the program strategies in a way that maximized their utility for the student). Others first consulted with trusted Pathways to Education program staff, informing staff of their reservations about the utility of the program content, which in turn allowed staff the opportunity to make personalized suggestions for how best to adapt the content rather than encouraging these participants to drop out of the program. In this way, Improvers creatively applied themselves in efforts to maximize their learning rather than passively accepting personally unfavorable circumstances.

Non-Improvers were less likely to adapt positively to get the most out of their program experience in the face of circumstances where they felt that the instructional content was not sufficiently challenging or stimulating. For example, Anwar, Nadia, Sandra and Sarah were some participants who made clear note of the fact that the program content was initially not sufficiently challenging for them and that they often felt "bored." However unlike Non-Improvers Sandra and Sarah, the Improvers Anwar and Nadia described certain actions that they took to mitigate the perceived disadvantage of being exposed to "repetitive" or previously learned content to make use of the intervention in order to meet their personal academic goals. Anwar described how he creatively modified the work he did during program sessions so that he could remain engaged and acquire new learning by bringing in his own academic texts for use in vocabulary instruction:

In the beginning it seemed a little too...it seemed very...how would you explain this, umm..i felt like I was past the level...so I told my SPSW...she told me...she wanted me to stay and she told me that if I had any homework or any other stuff that's related to English, I could be asking my tutor for help... the point of the program was to teach the students new vocabulary and stuff, so I used words from my homework and I learnt...basically I brought Philosophy homework, we'd read the text and ask my tutor what this means and she'd help. So I still learned stuff. I tried to take advantage of the time, and the tutor that was provided to me. (Anwar, Grade 11).

In contrast, Sandra, a Non-Improver, described her hesitation to inform her tutor of her personal learning needs, to the extent that she passively accepted her unmet needs: "Yeah, like I already like learned stuff like that, but I didn't want to like interfere with him and tell him like I don't want to do this, like I already know this, yea I just like went along" (Sandra, Grade 9). Improvers also demonstrated resourcefulness via their tendency to take actions to apply and consolidate their learning. For example Grace, an Improver, shared how she actively applied her learning in recreational contexts: "I learned some words like when I hear it on TV I'm like oh yea I learned that from the VLP program and I'm like so happy because I know what it means now" (Grace, Grade 9). Similarly, Anwar described how he made use of strategies such as KWL in the context of exam preparation:

Now I use it for studying for my exams, to study for it. Like I read through my notes, I know this I write it down, I want to know this, I write it down. So I use that and what I learned to see if I actually made improvement, so I'm actually sort of modifying it, but I'm also using it to my advantage. (Anwar, Grade 11).

Overall, although participants across both groups reported that they acquired new learning as a result of the intervention (e.g., new reading comprehension strategies, vocabulary), many noted

that various aspects of the program content represented a review of strategies learned in prior school grades. A key differential between the Improvers and Non-Improvers in this regard, was that the Improvers took agency to maximize their learning despite these circumstances – that is, they were resourceful in finding or creating opportunities to make the best use of their program time. They accomplished this in part by using their own academic texts for the purpose of program vocabulary instruction (i.e., for exposure to more complex vocabulary) and by purposively applying strategies on a consistent basis when engaged in recreational or academic reading activities.

Social (Peer and Family) Context

There were two themes at the social contextual level: reliance on peers and nature of parental support.

Theme 7: Reliance on peers

The extent to which participants relied on their peers either for supporting their engagement with the VLP or for support of their academic work in general, was a factor identified in this study to differentiate Improvers from Non-Improvers. On the whole, participants in this study shared that they made the decision to participate and remain in the intervention independently of familiar peer involvement (e.g., friends, similar-aged family members). However, Non-Improvers were more likely than Improvers to express that their behavioral engagement in the program (e.g., attendance, participation during sessions) was in part dependent on the co-participation of a friend or family member either within the same tutoring group or same session start and end time. Sandra, for example, shared that the participation of her "best friend" was critical to her own involvement, stating: "If she wasn't there I would definitely not have come." Priya (Non-Improver) also discussed the impact of her her sister and friend's co-involvement in the VLP: "I felt more comfortable if someone I knew was there, like someone to be there for you kinda." Improvers were less likely to be influenced by the involvement of familiar peers in the program, highlighting their own intrinsic motivation to participate. It is important to note that 80% of participants in this study had at least one friend or family member attending the VLP. Non-Improvers were also more likely to share that they often partly depended upon peers for homework support whereas Improvers were more likely to express that they were self-sufficient.

Theme 8: Nature of parental support

The quality of parental (and familial) motivation and encouragement as well as the extent of parental involvement in participants' academic activities differed among those who improved post-intervention and those who did not. Analysis of interviews with Improvers and Non-Improvers indicated that both groups of students received support and motivation from their parents and other family members to do well not only in the VLP but also more generally in school. Based on participant descriptions, parents of participants in both groups were significantly invested in their children's achievement based on numerous examples provided of the ways in which parents worked to encourage the participants to achieve. These students overwhelmingly came from families where education and hard work were valued, with parents purposefully motivating the participants to excel. There was a qualitative difference however, in the reported quality of parental or familial support provided to Improvers in comparison to Non-Improvers, which consequently impacted reasons why the participants were encouraged. Comparison of interview narratives of Improvers and Non-Improvers elucidated that both groups were motivated by their parents/families to achieve academically but for very different reasons. Improvers were motivated because they tended to: 1) absorb and internalize parents' verbal encouragement to work hard; 2) were consistently exposed to academically successful siblings or extended family who served as role models of achievement (e.g., siblings who completed postsecondary education, extended family members in traditionally prestigious professions such as

medicine, and/or 3) expressed aspirations to match the high achievement of model family

members or at least meet established familial high standards and expectations of academic

success. Anwar, for example, shared that he aspires to become a doctor because of the numerous

role models in his family:

Like everyone in my family, except for my dad, is like a doctor...I'm OK with that, I'm really happy about it. It's a little pressure I guess from my family, cause everyone is like a doctor, of any kind. (Anwar, Grade 11).

Improvers Grace and Nadia respectively share their motivations to succeed based on high

parental words of encouragement, high expectations and exposure to siblings in post-secondary

education:

They just like tell me that education is key, you need it...they expect me to be going to higher places like college and university, and they kind of motivate me, like when I have a good grade I'll show them and they'll be happy, "just continue the good work". Cause I'm the baby in the family, so they'll say we're expecting something from you in this family, so I do it... (Grace, Grade 9).

My mom motivates me a lot...she just talks to me a lot...even my older sister too cause now she's in University. She tells me to stay on my work and do all this stuff cause then it's gonna pay off cause slacking doesn't really help. (Nadia, Grade 9).

Non-Improvers in comparison, were motivated to achieve due to one or more of the following

reasons: 1) fear of "getting into trouble" with their parents should they underachieve; 2) parental

presence and monitoring of their behaviors in the school environment, and/or 3) the possibility

for financial or other external incentives (e.g., gifts, food) for compliance with their parents'

expressed desire for them to succeed. Sandra, for example, shared her parents' behaviors and

their impact on her motivation to remain in school:

Sandra: My mom's all up in my school...Like she goes everywhere. She's so like involved. A little bit too much, but yeah. She's always talking to my guidance counselor and everybody. And my dad just gives me money if I do good.Interviewer: Ok, so all that combined sort of makes you stay focused?Sandra: Yea...more like I'm afraid I'll get in trouble if I do bad.

Later in the interview, Sandra went on to express:

...without my parent's encouragement, I probably wouldn't care about school... Well I'd probably end up like my other friends. Like my friend.... it was like report card time right? But her mom doesn't even know she gets report cards all throughout the year. She only knows about the last report card at the end of the year [laughs]. Like, she doesn't know, she's not involved. And her dad's like MIA right? He's like sleeping or at work, so she never, she doesn't care! She cares, but she's not afraid of getting in trouble cause nobody's checking her homework. She's like "oh well". Like she tries but its not like "oh my God what am I gonna do when my parents get a phone call home". Nobody's home, you know, so she's not scared like that. But she still tries, but not as much as she would if her parents were on her back, you know? (Sandra, Grade 9)

In these ways, it can be argued that the quality of parental support, specifically the quality of parental efforts to motivate the participants, either supported participants' intrinsic motivation and responsibility for their own achievement (as observed in Improvers), or encouraged reliance on extrinsic reinforcement in a way that compromised personal responsibility (as observed in Non-Improvers). This can be further exemplified by Javitha, a Non-Improver (Grade 9) who expressed her dependence on food and monetary incentives offered by her mother in order to even attend school: "I probably wouldn't even go to school, if no one would force me."

The extent of parental involvement in participants' academic activities was another aspect of parental support that differentiated the two groups. Parents of Improvers not only consistently imparted motivating statements on their adolescents, but went a step further by taking a more hands-on approach to academic support provision. These parents would actively help with homework tasks, arrange for supplemental academic resources such as individual tutoring supports, or participate in parent-school organizations, for example. That is, these parents not only *said* motivating words, they *took action* to motivate their youth. Amala (Grade 9 Improver) shared her mother's active efforts to provide support: "My mom wanted me to go to a tutor... sometimes she tries arranging people to try to help me." Anwar, for example, described his father's role in helping with homework: "I turn to my dad, he seems to know everything kind of, it's weird. Every time I go for help, he seems to always be there for me, and he always helps me." Parents of Non-Improvers were described by these participants to be very willing to provide words of advice and encouragement and give youth verbal reminders to complete work, but were less likely to actively involve themselves beyond this. For example, Naya and Priya shared the extent of their parents' involvement:

They usually give me speeches and lectures, like that's all my parents do {giggles}. Like so I can keep myself on track and not get influenced by others and stuff like that. (Naya, Grade 11).

Whenever we have homework to do they just tell us to do it, they don't do anything much. (Priya, Grade 9).

Broader School, Community and Cultural Contexts

Theme 9: Perceptions of school

Participants' perceptions of their schools were an important contextual factor at the school level that differentiated Improvers and Non-Improvers. Specifically, the extent to which the self-reported curricular focus of participants' schools matched their personal goals and interests was a key factor that impacted their school perceptions. Improvers were more likely to endorse a match between the curriculum of the school they attend and their own individualized learning goals. Non-Improvers were more likely to express a mismatch between their learning interests and school curricular focus. Improvers frequently made comments suggesting that they were more academically-inclined in terms of their current and future goals, tending to express enjoyment in attending equally academically-focused schools, with strict teachers who push them to excel beyond their own conceived limits. Non improvers were more likely to complain about attending an academic school, often expressing desire for an alternative program more in line with personal interests (e.g. desire to take non-academic stream courses such as "cooking" or "steel pan"). For example, Sandra noted the following about her school: "It's like, they don't

have other options. Like for example [school Z] has a lot of other class options for extra classes, this school only has five things." Priya also shared her disdain that her school offers "Mostly academic stuff. Like other schools have cooking and more stuff like that but our school's mostly academic things."

Although all participants indicated that they generally liked their schools, Improvers were more broadly satisfied with their school experience and were better able to clearly describe factors that promoted or impacted their academic achievement, highlighting characteristics such as cultural and linguistic diversity, a school community accepting of individual differences, strict teaching, and an academically-focused curriculum. Anwar, for example, shared that he selected his school based on its provision of the International Baccalaureate (IB) academic program: "the one good thing about [school Y] is the IB program, and that's why I go there." Anwar also shared his feelings about the cultural and ethnic diversity at his school:

I feel that at school, I'm proud to be Bengali and Indian, and another good thing about my school is that there's no bullying of any sort, like no one picks on any other. Like I never notice...there's fights obviously between people, but there's never bullying, like exploding people or anything. And it helps everyone to be proud of who they are, and no one gets picked on because they are a specific race. Everyone's accepting at my school, and I feel that's one of the most important thing about schools, that everyone is accepted and no one is excluded or anything. (Anwar, Grade 11).

Nadia and Amala who attend the same school, described that they appreciated the fact that their school is "really diverse...a lot of immigrants" (Amala) and a "multicultural environment" (Nadia). Non-improvers equally commented on their appreciation of cultural diversity within their schools and positive perceptions of their teachers. However, they were more likely to also discuss negative aspects of their school experience (e.g., limited variety of course options, negative peer and teacher behaviors).

Theme 10: Engagement with school and broader community

Another factor that differentiated Improvers from Non-Improvers was their degree of engagement with their school community and broader community through involvement in cocurricular and/or extracurricular activities, as well as community volunteer work. Involvement in co-curricular or extracurricular activities was valued and desired by both groups of participants, however Improvers were more likely to be engaged in such opportunities. Many Improvers reported involvement in extracurricular activities that involved sports, arts, or academics. Improvers were also more likely to report current engagement in volunteer work in the community as well as participation in leadership training programs encouraging community leadership. As such Improvers appeared to be well-rounded and engaged citizens of their schools and broader community via their extracurricular involvements, unlike Non-Improvers who less frequently reported participating in extracurricular activities or community work.

Theme 11: Sense of cultural identity

One of the questions posed of participants in this study was regarding their cultural identity. Namely, participants were asked an open-ended question to explore the culture with which they identified or felt they shared similar values and beliefs. Improvers were more likely to exhibit a strong and clearly articulated cultural identity in comparison to Non-Improvers. All of the five Improvers expressed a strong identification with the culture espoused by their own or their parents' countries of origin. Grace, for example, shared her own identification with her Ghanaian origin:

I like my Ghanaian culture. I really enjoy it, cause that's how I was raised up. At school and stuff, I really embrace it with my friends or with my other friends, and when I go back to Ghana for vacation, sometimes I wish I was born there than I was born here. I just like it so much there and I really like my family back home and I just really like being Ghanaian. (Grace, Grade 9)

Similarly, Wajid shared that he identifies with his parent's culture, adding: "I barely know anything" about Canadian culture. Two of the five Improvers noted that they equally embrace a

Canadian cultural identity along with the culture of their country of origin. Non-improvers were more likely to express confusion, uncertainty, and/or conflicting feelings about their sense of cultural identity. While some Non-Improvers stated that their cultural identity was a "mix" of Canadian culture and that of their familial country of origin, others questioned the notion of the existence of any "pure" cultural identity altogether, and others, with expressed uncertainty, rejected the idea that they were anything other than Canadian.

Integrative Theme: Personal agency

Integration of themes resulted in one theme that spanned across the four major contextual areas and related the 11 primary themes -- personal agency. Through an examination of the various themes in this study thus far, it became apparent that among the Improvers, there was an aspect of active effort to achieve despite adverse circumstances. Both Improvers and Non-Improvers were Grade 9 to 11 students attending the VLP intervention within the Regent Park Community in Toronto. Based on the neighborhood context and other socio-demographic variables indicated by prior research to place such students at risk for academic underachievement, all of the participants in this study were effectively considered to be at risk for low achievement and school disengagement. However, it was very clear that there were various protective factors at play which mitigated the risk for the majority of these students given their collective expressions of an ongoing desire to achieve and make meaningful contributions to society, irrespective of intervention outcome status. Comparison of contextual factors (both risk and protective) at play among Improvers and Non-Improvers following this reading intervention, helped to elucidate additional broader factors that potentially served to even further mitigate the at-risk status of Improvers, providing clues about the kinds of enduring factors that can make a difference in minimizing academic risk within different ecological contexts of influence.

Through an examination of the various themes in this study thus far, it became apparent that among the Improvers, there was an aspect of *active* effort to achieve despite immediate circumstances. That is, these students exhibited a strong sense of personal agency, defined here as the capacity and propensity to take purposeful action in exerting influence over one's own environment. At all contextual levels, the identified themes could be related based on their exemplification of the broader concept of personal agency- the capability to actively intervene and influence one's learning environment and pathways of skill achievement; in this case, vocabulary and reading comprehension skill development. At the individual level, Improvers were the students who were achievement-oriented, had a positive academic self-concept and a stable sense of their future goals. The commonality is that these students took action in seeking out necessary supports to help them achieve, refusing to settle for insufficient supports. They were self-efficacious and self-assured, unlike the Non-Improvers. Within the program context, Improvers were found to be more actively engaged with the program content, applying and making use of newly acquired knowledge. They were more willing to make adaptations to their program experience in order to maximize their learning, and independently sought out resources to help themselves, again exhibiting their role as active agents in the learning process. At the social contextual level, parents of Improvers also embodied personal agency in that they directly supported and encouraged their youth to achieve via helping them directly with homework, exposing them to academically successful role models and effectively taking a hands-on approach to making a meaningful difference in their children's learning trajectory. It is conjectured that this parental agency thus encouraged more intrinsic motivation in the youth, who were more inspired to take responsibility and personal agency for their own development. At the broadest contextual level, this theme of personal agency was predominant, exemplified by the Improvers' actions in purposefully engaging with their school and neighborhood through

extracurricular and volunteer work. It was also demonstrated by the Non-Improvers' willingness to passively remain in academic programs that were not suited to their personal goals or interests. In summary, integration of themes indicated that personal agency was a critical differentiating factor between both participant groups and across all contexts. A high degree of personal agency potentially served as a protective factor that mitigated the likelihood of poor outcomes following the preventive reading intervention.

3.4 Summary and Interpretation of Findings

Using an ecological framework as a guide for inquiry, the current study examined the factors differentiating adolescents whose reading comprehension outcomes improved or did not improve following an after-school reading intervention (the Vocabulary Learning Project, or VLP). The systematic synthesis of data from in-depth interviews with participants resulted in the identification of a variety of individual and broader contextual factors (indicated by thematic findings) potentially associated with differential reading outcomes following participation in the VLP.

Differential factors at the Individual Level

At the individual level, the factors that played a role in differentiating program Improvers and Non-Improvers included achievement orientation, academic self-concept, and sense of future aspirations. Existing literature regarding the role of these factors on achievement highlights the critical role of motivation as the underlying common factor implicated, as will be discussed shortly.

Improvers and Non-Improvers were generally inclined to do well academically, consistent with recent data indicating that immigrant and language minority students in Canada are 30% less likely than native students to report low achievement motivation (OECD, 2018b). Achievement orientation in this study is synonymous with Dweck's (1986) concept of achievement motivation and may help explain the differential performance between the two groups of participants in this study, as it indicates that certain types of achievement goals lead to certain behavioral patterns. Achievement motivation that is mastery-oriented (focused on improving competence) or performance oriented (focused on demonstrating competence) has been shown to be related to engagement in more adaptive, effortful and tenacious academic behaviors, which then positively impacts achievement (Elliot, McGregor & Gable, 1999; Nazarieh, 2015). Improvers likely differed from Non-Improvers in this regard as they reported engagement in these behaviors.

Improvers reported higher academic self-concept. Higher academic self-concept is associated with increased academic achievement and this relationship is reciprocal and mutually reinforcing (Marsh, 2007). It has also been established that adaptive motivational beliefs promote achievement (Bandura, 1997; Wigfield & Eccles, 2000). Higher achievement also fuels increased achievement orientation, thus findings of this study suggest that these factors of greater achievement, higher academic self-concept and achievement goal orientation operate in a cyclical reinforcing manner, with heightened motivation underlying these relationships. In fact, research has shown that academic self-concept is highly associated with motivation (Bong & Skaalvik, 2003) as well as achievement orientation (Nazarieh, 2015). Motivation has been also linked to career and educational aspirations (Bandura, Barbaranelli, Caprara, & Pastorelli, 2001). Academic self-concept has considerable and direct effects on academic aspirations (Marsh, 1991) through its interrelationship with academic engagement (Bandura et al., 2001; Wigfield & Eccles, 2002). Taken together, the current study findings suggest that the mutually reinforcing cycle involving academic self-concept, achievement orientation and achievement likely contributes to adolescents' consequent likelihood to prepare and plan for future career

aspirations, consistent with Improvers' greater tendency to report having defined career goals. Therefore, to understand the contextual factors (or risk and protective factors) at the individual level that differentiate intervention responders from non-responders in this culturally and linguistically diverse sample beyond cognitive and linguistic factors, I argue that academic selfconcept operates in tandem with motivation, together playing a "complementary and synergic" role in impacting academic trajectories (Green et al., 2012, p. 1120) and future career prospects by their direct impact on academic engagement. Engagement reflects the manifestation of motivation and self-related beliefs (i.e., academic self-concept). In fact, engagement has been shown to be a mediator of the relationship between academic self-concept and achievement (e.g., Bong & Skaalvik, 2003), so present study results suggesting decreased engagement on the part of Non-Improvers are likely connected to those individual level factors of self-concept and consequent motivation (Skinner, Kindermann, Connell, & Wellborn, 2009).

Differential Factors at the Intervention Program Level

Within the context of the intervention, the factors that played a role in differentiating program Improvers and Non-Improvers included participants' personal investment in positive outcomes, participants' resourcefulness and perceptions of the tutoring group. The themes of personal investment in positive outcomes and resourcefulness highlight at their core the participants' varying engagement with the program, a by-product of their motivational differences at the individual level as previously highlighted. Non-Improvers exemplified their weaker motivation (in comparison to the Improvers) by reporting less adaptive strategies when faced with similar tutoring environmental circumstances as Improvers (e.g., dislike of program content, perceiving the content to be insufficiently challenging). This was exemplified by lessened investment of Non-Improvers in positive outcomes, decreased resourcefulness as well as increased vulnerability to adverse tutoring group factors (e.g., experiences of diminished

program interest due to peers' minimal engagement). Overall, students' individual level characteristics interacted with the intervention context such that they were less likely to be engaged (or at most displayed tenuous engagement). Social contexts play an important role in influencing intervention outcomes, as will be further elaborated by an interpretation of findings at the social contextual level.

Differential factors at the social contextual level

In this study, Improvers and Non-Improvers were differentiated by the extent of influence of peer relationships and the nature of parental (and familial) support offered to them. Those students whose outcomes improved post-intervention were less vulnerable to negative peer influences and had parents who provided them with both encouragement and instrumental academic support. Existing literature indicates that adolescents' social contexts can positively influence their pursuit of goals in general, if there is correspondence between personal goals and those of others in the environment (Wentzel, 2005) such as peers and parents. Conversely, youth can be quite negatively impacted by social contexts where their peers do not actively promote the pursuit of academic goals (Wentzel, 2005; Ryan, 2000). Well-established literature suggests that peer relationships in learning contexts play a critical role in promoting school engagement and achievement (Fredericks et al., 2004). In fact, existing literature suggests that children and adolescents make efforts to select friends who are similar to them on dimensions such as academic achievement, attitudes and interests, among others (Bagwell & Schmidt, 2011; Schneider, Wiener & Murphy, 1994). The findings of this study indicate that differential performance of participants following intervention may not only be related to differences in perceived peer relationships (consistent with past findings). Differential performance may also be accounted for by the interaction of these contextual factors with individual characteristics such as motivation and engagement, and can make the critical difference between having

positive and negative outcomes once these ecological contexts are jointly considered within and across youth with otherwise similar background characteristics.

It is also possible that the influence of youth's parents and families on their academic goal pursuits and by extension on their intervention program outcomes (as described by participants) may be a result of parents' more direct impact on youth motivation and engagement.

Consistent with prior research findings that young people from lower socioeconomic backgrounds often view education as the means to avoid the difficult lives of their parents and improve their own lives (Lopez, 2001), the majority of participants in this study highly valued their education. However research also indicates that parents of cultural and ethnic minority youth are often able to provide emotional support but not the vital instrumental support the youth need to mobilize educational excellence (Dennis, Phinney & Chuateco, 2005). Although collectively parents of both groups of participants reportedly held strong family values toward education -- a finding that has been supported by other research in immigrant populations (OECD, 2018b) - Improvers' parents provided more instrumental support, exposed their youth to familial exemplars or role models of achievement as well as instilled in them the expectation of post-secondary education. These are protective factors that have been found by other researchers to be associated with academic engagement (e.g., Roundfield et al., 2018) and that can counteract the academic risk factors youth face (Dennis et al., 2005). Such parental behaviors were potentially facilitative of intrinsic sources of motivation in youth due to lessened reliance on external reinforcers of achievement (e.g., compensation, gifts, threat of punishment). These external reinforcers can have the unfortunate effect of encouraging less enduring extrinsic motivation or self-directive behavior that is transferrable to other learning contexts like the VLP, as exemplified by parents of Non-Improvers. Improvers' parents (in comparison to parents of

Non-Improvers) were able to provide more adaptive supports to their children for a variety of potential reasons that were beyond the scope of investigation in this study, such as: higher parental education; better familiarity with the Canadian school system; better English language skills; more physical availability to provide direct support, and more cultural capital, to name a few. Follow up integration of qualitative findings from this study with quantitative measures from Study 1 may prove helpful for a more thorough exploration of these group differences.

Differential factors within the broader environment (school, community and cultural aspects)

Youth's perceptions of their schools, engagement with their schools and community, and their sense of cultural identity were the broader contextual factors found to differentiate Improvers from Non-Improvers in this study. It has been shown that school perceptions influence personal goals (Greene, Miller, Crowson, Duke, & Akey, 2004), consistent with reports of Non-Improvers in this study who were more likely to describe attending schools where the curricular focus did not support their personal learning goals or interests. Involvement in extracurricular activities has been generally linked to increased engagement in the school environment (Mahoney & Cairns, 1997). Consistent with this finding Improvers were more involved in extracurricular activities at school as well as engaged in community volunteer work, compared to their counterparts who did not respond to the intervention. In being more engaged, Improvers likely represent youth who regularly experience a deeper sense of community connectedness and belonging, which is in turn potentially related to increased motivation. This interpretation is reasonable in light of prior research findings that the relationship of academic achievement with academic motivation can be mediated by affective and behavioral engagement (e.g., voluntary participation in school-based activities, positive appraisals of school and academic subjects) (Green et al., 2012). In fact, it has been found in prior research that among urban adolescents at risk of high school dropout, adolescents' sense of belonging at school is significantly associated

with motivation (Goodenow & Grady, 1993). A sense of belonging likely matters considerably for these participants given that adolescence is a life stage focused on the crucial developmental task of personal identity formation. The capacity to form such an identity is filtered through social experiences, relationships and perceived group belonging (Guruge & Collins, 2008). Community belongingness is particularly critical when an adolescent is also a first or second generation immigrant living in a culturally and linguistically diverse community. The importance of community belonging was reflected in the influence of participants' perceived cultural identification as a differentiating factor between the groups: those who fared better postintervention were more likely to express a strong and stable sense of ethnically-based cultural belonging with either their culture of origin or Canadian culture. Conversely, Non-Improvers were more tentative or undecided about their cultural identity. Interestingly, it has been found that adolescent students' association with different cultural and ethnic groups is an important aspect of social relating that influences motivation (Goodenow & Grady, 1993). Thus for Non-Improvers, uncertainty around their cultural identity may have a negative impact on their engagement and learning at least through its impact on motivation. An implication of this finding is that the design of similar reading interventions may benefit from the incorporation of culturally relevant approaches so that students have the opportunity to see themselves reflected in program curricula.

The overarching roles of agency and motivation

The various thematic findings of this study have demonstrated the critical role of differences in motivation as a core differential factor between intervention responders and non-responders. Motivation is the factor that likely drove the demonstration of personal agency (as previously defined) among participants in this study, ultimately influencing the likelihood of intervention success or failure. Improvers were active agents in the learning process,

purposefully intervening and influencing their own learning environment and pathways. Research has shown that students with agency have a strong sense of self efficacy; they feel empowered to take action to their accomplish goals (Johnston, 2004). Students who take agency over their learning are more motivated and consequently more likely to be successful academically (Lin-Siegler, Dweck & Cohen, 2016).

Resilience research, particularly that involving socio-economically disadvantaged children and youth, often postulates that "individuals vary in their ability to overcome disadvantage because of their willingness and ability to mobilise their own psychological and physical resources, and the resources available in their social and physical environment" (OECD, 2018b, p. 33). Essentially this indicates that a critical marker of resilience is the capacity and propensity to take personal action (i.e., personal agency) fueled by motivation. If it can be assumed that both groups of youth in this study share similar background characteristics and some similar contextual circumstances yet responded differently to the same reading intervention, it can be extrapolated that factors differentiating Improvers from Non-Improvers are indicative of potential risk and resilience factors. Overall, integration of thematic findings in this study indicate that motivation and consequently, personal agency, were key resilience factors that were associated with improved reading comprehension outcomes in response to an afterschool reading intervention among disadvantaged culturally and linguistically diverse youth. The varying motivational patterns of participants in either group were manifested in different ways across all contextual areas under study. The concept of resilience as it applies to this study essentially posits that youth who are living in vulnerable environmental contexts characterized by social and economic adversities can regardless thrive (e.g., experience improved reading outcomes) in the presence of protective factors, which when identified can inform the amelioration of interventions to maximize student outcomes.

Study Limitations and Future Directions

There were several notable limitations to the present study. The findings were limited by the single measure used to demarcate responsiveness or non-responsiveness to the intervention. Certain demographic characteristics that may be associated with response to intervention were not explored (e.g., gender, country of origin, learning problems such as learning disability, neurodevelopmental disorders such as Attention-Deficit Hyperactivity Disorder (ADHD)). For example, differential levels of personal agency among Improvers and Non-Improvers may be partially attributable to learning differences and differences in cognitive functioning, such as individual differences in working memory, attention, and executive functioning. The presence of ADHD and associated neurocognitive deficits, particularly deficits in executive functions such as initiation, planning, sustained effort, and self-monitoring may relate to differential levels of agency (Wiener & Daniels, 2016). Future studies that take into account such factors as intraindividual learning and cognitive profiles will support a more comprehensive understanding of factors differentiating Improvers from Non-Improvers. Future studies may also take advantage of a larger sample size and multiple comparison groups of cases based on these other characteristics, as the use of multiple comparison groups permits for a more nuanced understanding of key differentiating factors. It may be beneficial to explore differentiating factors quantitatively, for example by using findings from this study to help identify relevant quantitative measures for theoretical modelling. Predictor measures of the various contextual factors identified in this study can then be analyzed quantitatively to test whether they are predictive of specific outcomes or assess their role as moderators or mediators of outcomes. Additionally, longitudinal exploration of outcomes may help identify contextual factors differentiating adolescents with differing growth profiles, for example. Despite its limitations, the current study was a theory driven qualitative exploration based on a multidimensional

ecological model of factors proximally and distally related to differential reading comprehension skill development post intervention among culturally and linguistically diverse youth in a socioeconomically disadvantaged community.

Potential Implications of Findings

A growing body of literature emphasizes the importance of studying successful and resilient youth in order to promote more positive youth skill development (Larson, 2000). Findings from this study provide opportunities for the design of future interventions, particularly given that academic motivation is not a fixed trait, but is rather malleable and open to intervention (Green et al., 2012). Reading interventions that thus prioritize improved motivational outcomes as well as the other key reading outcomes may be well positioned to expect increased engagement and subsequently more positive intervention effects. Other researchers have noted the importance that interventions, and that evidence-based practice means that evidence-based interventions must be tailored to individual learners (e.g., Kratchowill, 2003). These issues may be especially important for adolescent language minority and disadvantaged youth, where research is very limited and it is yet unclear which practices have optimal chances of improving outcomes (Snyder et al., 2017).

Chapter 4 – A Mixed Methods Evaluation of the Vocabulary Learning Project (VLP): Integration of Findings from Studies 1 and 2 (Study 3)

4.1 Introduction

Researchers have been increasing their efforts to develop and evaluate reading interventions targeted at immigrant and language minority adolescents. Variability in individuals' response to intervention can contribute to smaller intervention effects (Fletcher & Vaughn, 2009). Integration of quantitative and qualitative findings can be instrumental in elucidating broader contextual issues related to differential intervention efficacy and inform overall intervention utility. To date, there are a significant dearth of experimental studies evaluating reading interventions among culturally and linguistically diverse youth, particularly in the Canadian context.

This study represents the third and final phase of a multi-phase, quasi-experimental mixed method evaluation of a reading intervention that targeted the development of adolescent vocabulary and reading comprehension skills. The overarching aim of the program evaluation was to identify factors that are associated with differential intervention effects at the individual level (Study 1) and within multiple environmental and individual contexts (Study 2). In the first phase of the research (Study 1), the specific aim was to establish the effectiveness of the intervention, identify subgroups of participants for whom the intervention may be especially effective, as well as explore person-level predictors of gains in reading comprehension skills following intervention. In the second phase of the research (Study 2), person-level and contextual factors potentially associated with differential response to the intervention were explored. The final phase of the research (e.g., the present study) involves an integration of findings from phases one and two, summarizing the overall potential of this intervention as a tool for reading skill improvement.

165

Across the first two phases of this research, the primary researcher used a combination of qualitative and quantitative methods. Quantitative methodology was used in Study 1 to assess treatment effects, differential program effects (via subgroup analyses), and predictors of change. Qualitative methodology was used in Study 2 to explore variability in response to the preventive reading intervention, as well as to examine contextual factors potentially implicated in differentiating poor program responders (i.e., *Non-Improvers*) from good responders (i.e., *Improvers*).

4.1.1 **Purpose of the current study**

The ultimate aim of this study was to gain an ecological profile of factors implicated in differential response to the VLP reading intervention based on findings from integration of quantitative (Study1) and qualitative (Study 2) data. The secondary aim of the study was to better understand the utility of this intervention in maximizing the reading skill development of immigrant and language-minority youth (e.g., discover which intervention design modifications are most likely to improve the likelihood of effectiveness for a greater proportion of participants). Thus, findings from this study will enable an exploration of the implications for the design of future interventions in culturally and linguistically diverse contexts.

4.1.2 Research Question

The following mixed method research question was posed for the present study:

(1) How do findings from the qualitative study enhance findings from the quantitative

study to facilitate understanding of the differential impact of the VLP intervention? The present study will address this overarching mixed method research question by comparing data from the larger impact study with multiple case study data from the qualitative study based on a subsample of cases who either demonstrated improvement or non-improvement in reading comprehension skills at post-intervention testing. It was hypothesized that the qualitative themes combined with quantitative findings would help inform an integrated understanding of the effectiveness of this intervention and the contextual factors which influence its effectiveness.

4.2 Methods

4.2.1 Research Design: Mixed Methods Approach

This study involved a sequential embedded quasi-experimental mixed methods design consisting of three distinct yet interrelated phases (Creswell & Plano-Clark, 2007). In the first phase, three sequential cohorts of youth were recruited to participate in the experimental study over three years from Regent Park's Pathways to Education community program (see Chapter 2 for a description). Each sequential cohort was implemented as pretest-intervention-post-test. All participants were tested on a battery of measures of the study outcomes (i.e., reading comprehension and vocabulary knowledge) and potential predictors (e.g., motivation for reading) at pre- and post-intervention. Demographic information was also gathered. Program fidelity was also monitored in phase one by the concurrent collection of observational data of every intervention session through audio-recordings of entire sessions. As previously stated, the aim of this phase of the study was to compare the intervention and control group participants' post-test scores with respect to the intervention outcomes, in order to ascertain the impact of the reading intervention (i.e., to determine whether the intervention was effective in improving reading comprehension and vocabulary knowledge). Another goal of the initial study phase was to determine whether certain person-level factors (demographic, linguistic and psychological) assessed pre-intervention, were predictive of response to intervention. This partially addressed the study's secondary objective (i.e., determining person-level factors associated with program outcomes that were measured quantitatively). Lastly, the first phase informed program

implementation (e.g., process, fidelity, student engagement) through observational data gathered from audio-recordings.

The second phase of the study also addressed the study's secondary objective by allowing for inductive exploration of the person-level and broader contextual factors that may influence program impact. In this phase, participants were purposefully selected to provide in-depth interview accounts of their experiences in the intervention, based on their individualized program outcomes (as determined by results of the first phase, which indicated whether each students' outcome scores improved or did not improve post-intervention). The goal was thus to determine whether there were differences in the thematic patterns emerging from interviews with intervention non-responders compared to interviews with intervention responders, and ultimately identify factors differentially associated with youths' response to intervention. This embedded quasi-experimental design allowed for the inclusion of a qualitative strand of inquiry that addressed a secondary research objective within a broader program evaluation objective. The qualitative strand was sequentially embedded to allow for elaboration and enhanced interpretation of the findings from the primary impact study. In particular the qualitative phase allowed for examination of the multi-level contexts potentially associated with differential response to intervention, through exploration of the experiences of participants with certain kinds of outcomes (Creswell & Plano-Clark, 2007). In this design, quantitative and qualitative data interfaced at the design level. The two strands of the study interacted with each other, with the results of the qualitative strand dependent on the preceding quantitative strand. The qualitative and quantitative strands were equally weighted in order to assess both program impact and process/context respectively, to maximally account for the potential of the program.

This design was especially appropriate for evaluating a community-based reading intervention in a multicultural context because of the need to capture both direct program impact
as well as the interpersonal and contextual factors at multiple ecological levels that influenced the intervention's effectiveness (e.g., parental support, personal investment in positive outcomes). The quantitative and qualitative data sources were integrated in the current study (final phase) which sought to achieve coherence across all of the study's research questions in order to address the mixed method research question. The overall study summary and mixed methods model are depicted in Table 22 below and Figure 1 (from Study 1) respectively.

Table 22

Overall Mixed Method Study Summary

Research Question	Sample	Data Sources	Analytic Method
What is the impact of an intensive after-school reading intervention on the vocabulary knowledge and reading comprehension skills of academically at-risk youth when compared to a business as usual control condition?	Youth recruited from Pathways to Education - Regent Park site; some assigned to intervention group, others assigned to control	Test battery administered at pre- and post-intervention that includes standardized measures of the outcomes (reading comprehension, vocabulary knowledge)	Pre-post analyses of participant outcomes (program impact) via mixed ANOVA
What is the impact of the intervention on the following baseline (pretest) subgroups of students relative to their control group counterparts:1) students who are struggling (i.e., below average) with their language comprehension skills, and students who are not struggling (i.e., at least average), 2) students who report low levels of reading strategy use and students who report at least average reading strategy use, and 3) students who report at least average motivation to read and students who report below average motivation to read?	Intervention group participants	Test battery administered at pre- and post-intervention that includes standardized measures of the outcomes	Pre-post analyses of participant outcomes (program impact) for subgroups of participants via mixed ANOVA
Do the following person-level student factors differentially predict response to intervention after accounting for length of residence (in years) in Canada and nonverbal ability: a) baseline language comprehension skills; b) motivation for reading and c) reading strategy use?	Intervention group participants -	Test battery administered at pre- and post-intervention that includes standardized measures of the outcomes and identified predictors; demographic information gathered at pre-test	Hierarchical linear regression

What contextual factors are uniquely associated	Purposefully	Semi-structured	Within and cross-case
with responsiveness and/or non- responsiveness	selected	interview that addresses	thematic analysis
to the intervention from the perspective of	subsample of	youths' broader	
participants?	intervention	contextual experiences	
	group	as program participants	
	participants,		
	each categorized		
	based on their		
	outcome data		
	from Q1 as		
	either program		
	Improvers or		
	program Non-		
	Improvers		
How do findings from the qualitative study	Connect the	Findings contributed by	QUAN outcome data
enhance findings from the quantitative study to	samples from	select measures used in	linked to QUAL case
facilitate understanding of the differential	questions 1-3	questions 1-3	interviews; develop within-
impact of the VLP intervention?			case narratives and cross-
			case themes that are
			linked to specific program
			outcomes; interpret QUAL
			data in context of QUAN
			(including predictors from
			Q2); create summary
			interpretation of program
			impact with QUAL data
			serving to elaborate
			factors that contribute to
			program impact

4.2.2 Data Analysis

Within a sequential embedded quasi-experimental mixed method research (MMR) design, there are three primary data analysis steps: a) analysis of primary data (e.g., quantitative data), b) analysis of secondary data (e.g., qualitative data) followed by c) integration of quantitative and qualitative data to determine how secondary data augment primary data (Creswell & Plano Clark, 2011). There are a variety of data integration strategies for MMR. One strategy for comparing results involves side-by-side comparisons in a discussion or summary table. A further application of this method, particularly when qualitative data is derived from a small selection of cases, involves the creation of a mixed methods matrix (O'Cathain et al., 2010; Miles & Huberman, 1994; Creswell & Plano Clark, 2011).

A mixed methods matrix was created for data integration. This matrix essentially represented a table of case profiles from the ten purposively selected cases in Study 2. Characteristics of each case, using key variables from measures in Study 1 as well as cross-case thematic findings from Study 2 were summarized. *Improvers'* profiles were constantly compared to Non-Improvers' profiles by examining areas of similarity and/or divergence per variable. This allowed for a visual display and analysis of the manner in which the qualitative and quantitative findings interfaced such that cross-methodological patterns could be illuminated. The relevant variables within the matrix were organized such that they combined to form an ecological profile across participants that necessarily transcended individual contexts to broader contexts experienced by participants. This approach of creating an ecological profile is consistent with Bronfenbrenner (1979, 2005)'s ecological theory of human development. More specifically with respect to reading, it is consistent with the expanded SVR model (which is also based on ecological theory). The ecological profile approach served as a unifying theoretical framework for which the qualitative and quantitative research strands could offer different types of evidence and understanding. Such a theoretically grounded integrative approach helps achieve coherence across the research questions when qualitative and quantitative data sources are integrated (Creswell & Plano Clark, 2011). Building each case's profile using the methods matrix helped to identify areas of similarity and difference that then pointed to the ways in which the qualitative strand of research extended the knowledge gained in the quantitative strand.

4.2.3 Measures

Creating a mixed methods matrix (or, profile table) involved visually linking relevant variables from Study 1 to qualitative themes from Study 2. Table 23 summarizes the measures

from Studies 1 and 2 selected for incorporation into the matrix. Specific descriptions of the

measures were provided in Studies 1 (Chapter 2) and 2 (Chapter 3).

 Table 23

 Measures Used for Integration of Findings

 Measure

 Demographic

 Family Background Questionnaire

 Pre-Interview Questionnaire

 Cognitive

 Matrix Analogies Test

 Linguistic, Psychological and Behavioral

 Language Comprehension composite

 Motivation for Reading Questionnaire

 Metacognitive Awareness of Reading Strategies Questionnaire

 Broader Contextual

 Semi-structured Interview Guide

4.3 Results

The methods matrix is displayed in Table 24. This section summarizes the results obtained from an analysis of the matrix at each ecological area of interest. Characteristics within each ecological contextual area were compared across *Improvers* (n=5) and *Non-Improvers* (n=5).

Table 24

Methods Matrix: Profile of Case Characteristics for Mixed Method Integration

	IMPROVERS						NON-IMPROVERS					
	Wajid	Amala	Nadia	Anwar	Grace	Naya	Sandra	Javitha	Priya	Sarah		
INDIVIDUAL CONTEXT												
Demographic												
Gender	Male	Female	Female	Male	Female	Female	Female	Female	Female	Female		
Age	14	14	15	16	15	16	15	14	14	15		
Grade	9	9	9	11 Bang-	9	11 Bang-	9	9	9	9		
Country of Birth	Canada	Canada	Canada	ladesh	Canada	ladesh	Canada	Canada	Canada	Canada		
Age at Immigration	N/A	N/A	N/A	9	N/A	7	N/A	N/A	N/A	N/A		
Length of Residence in Canada	14	14 Bengali,	15	7	15	8	15	14	14	15		
First Language(s)	Tamil	Urdu, Hindi	English	Bengali	Twi	Bengali	Harare/ Amharic	Tamil	English	English		
Second Language	English	English	Arabic	English	English	English	English Harare/	English	No L2	No L2		
Language(s) spoken at home	Tamil	Bengali, English	English, Arabic	Bengali	Twi, English	Bengali	Amharic, English	Tamil, English	English	English		
Cultural Background	Asian	Asian	African	Asian	African	Asian	African	Asian	Bicultural	Bicultural		
Cognitive												
Non-verbal ability Linguistic, Psychological and Behavioral	106	87	91	92	88	88	91	82	95	84		
Language Comprehension Level	Good	Good	Good	Good	Good	Poor	Good	Good	Good	Good		
Motivation for Reading Level	Good	Good	Good	Good	Good	Poor	Good	Good	Good	Good		
Reading Strategy Use Level	Good	Good	Poor	Good	Good	Poor	Good	Good	Good	Good		
<i>Contextual Themes</i> Theme 1: Achievement Orientation	Improver achievem	<i>Improvers</i> more oriented towards academic achievement <i>Improvers</i> exhibited more positive academic self- concept					Non-Improvers less likely to report achievement-seeking					
Theme 2: Academic Self-Concept Theme 3: Sense of Future	concept						Non-Improvers exhibited less positive academic self-concept					
Aspirations	Improver	s readily con	nveyed vario	ous future asp	pirations	Non-Improvers did not readily convey future aspirations						
PROGRAM CONTEXT												
# VLP Sessions completed	13	9	11	12	10	13	10	11	13	13		
<i>Contextual Themes</i> Theme 4: Investment in Positive VLP Outcomes Theme 5: Tutoring Group Climate Theme 6: Participant Resourcefulness	<i>Improvers</i> more invested in having positive program outcomes <i>Improvers</i> perceived tutoring group as positive and suportive <i>Improvers</i> were resourceful in finding ways to maximize personal program benefit					<i>Non-Improvers</i> less invested in personal program outcomes <i>Non-Improvers</i> perceived tutoring group as less supportive <i>Non-Improvers</i> more likely to passively accept unfavorable program experiences; less resourceful						
		S	OCIAL CO	NTEXT: PE	EERS AND	FAMILY	high					
Maternal Highest Education Level	college diploma	college diploma	college diploma	university degrees	school diploma	high school	school diploma	high school diploma	college diploma	college diploma		

Determal Hickory Education Laura	college	college	college	multiple university	college	high school	college	high school	less than high	less than high	
Paternal Highest Education Level	dipionia	dipionia	dipiona	degrees	uipionia	dipiona	dipiona	dipiona	school	school	
Contextual Themes											
	<i>Improvers</i> less influenced by participation of peers; <i>Non-Improvers</i> were influenced by participation of peers;									ers;	
Theme 7: Reliance on Peers	personally	driven to p	participate in	n VLP		socially driven to participate in VLP					
Theme 8: Nature of Parental	Improvers' parents are more actively supportive of					Non-Improvers' parents were more passively supportive of					
Support	youth aca	demic ende	avors			youth academic endeavors					
BROADER SCHOOL, COMMUNITY & CULTURAL CONTEXTS											
Contextual Themes											
	Improvers were more likely to have positive Non-Improvers were more likely to highligh							ght unfavore	d aspects		
Theme 9: Perceptions of School	perceptions of their school environment of their school environment										
Theme 10: Engagement with	<i>Improvers</i> were more engaged with their school and					Non-Improvers were generally less engaged with school and					
School & Community	neighborhood (e.g., volunteering, extracurriculars) communiity										
Theme 11: Sense of Cultural	Improvers	s exhibited a	a more stabl	e and solidly		Non-Improvers exhibited confusion, uncertainty and conflic					
Identity	developed	l sense of cu	ltural identi	ity		feelings about	ut their sense	of cultural identi	ity		

Individual Level Context: Demographic Factors

Given that participants in both groups were matched by age and nonverbal ability, these two variables were only included descriptively as they were controlled for in advance. Although there were no male Non-Improvers, analyses from Study 2 indicated no significant gender differences among participants in either outcome group. Eighty percent of participants from each group were Canadian-born. *Improvers* and *Non-Improvers* were also similar in terms of average length of residence in Canada (M=13, *SD*=3.39 and M=13.2, *SD*= 2.95 respectively). Participants in both groups exhibited diversity in first languages spoken. Only one *Improver* reported English as a first language (i.e., native English speaker or non-language minority); two *Non-Improvers* reported that they were native speakers. For the four *Improvers* whose first language was not English, English was in fact the second language. Of the three *Non-Improvers* whose first language other than English at home, while 60% of *Improvers* spoke a home language other than English. Both groups displayed comparable diversity with respect to cultural background.

Individual Level Context: Linguistic, Psychological and Behavioral Factors

Both groups' profiles of language comprehension, motivation to read and reading strategy use were comparable; that is, these variables did not differentiate *Improvers* from *Non-Improvers*. Based on their baseline scores on these variables, All *Improvers* and *Non-Improvers* fell within the range of "good" language comprehender and "good" motivation to read at pre-test, with the exception of one *Non-Improver* who fell in the "poor" range on these variables as well as on reading strategy use. Only one *Improver* fell in the "poor" range with respect to reading strategy use; all others reported "good" use of reading strategies.

Individual Level Context: Thematic Findings

A thorough discussion of the thematic findings and the manner in which themes represented individual-level contextual areas that differentiated the two groups can be found in Study 2. Generally, *Improvers* were more oriented towards academic achievement, exhibited more positive academic self-concept and readily conveyed various future career aspirations, relative to *Non-Improvers*. Although *Non-Improvers* described that they valued academics, they were less likely to report being high achievement seekers, exhibited a less positive academic selfconcept, and did not readily convey future aspirations.

VLP Program Level Context: Program Attendance

In terms of the proportion of VLP program sessions completed, *Improvers* completed a median of 10.5 sessions (81%), while *Non-Improvers* completed a median of 12 sessions (92%). *VLP Program Level Context: Contextual Themes*

A thorough discussion of the thematic findings and the manner in which themes represented program-level contextual areas that differentiated the two groups can be found in Study 2. Compared to *Non-Improvers*, overall *Improvers* were more affectively and behaviorally invested in positive program outcomes (i.e., improving vocabulary and reading skills), perceived their tutoring groups as positive and supportive of their personal program goals, and were generally more resourceful in finding ways to maximize individual program benefits (e.g., seeking out more challenging reading material during sessions, encouraging their tutor to adapt program material to personal learning needs).

Peer and Family Social Level Context: Parental Education

Parental education, although broadly a demographic variable, was in this study considered to be a factor more relevant for inclusion in the peer and family level context given its reflection of family-related issues. It has also been well established based on prior literature that children and youth's academic outcomes can be moderated by parental education, particularly among socio-economically disadvantaged populations. The proportion of parents who reported having completed higher levels of education (e.g., college diploma or university degree) across both groups were comparable (100% of *Improvers* ' parents and 80% of *Non-Improvers* ' parents). However, while every *Non-Improver* had at least one parent whose highest education level was a high school diploma, only one *Improver* had a parent for whom high school was the highest level of education.

Peer and Family Social Level Context: Contextual Themes

A thorough discussion of the thematic findings and the manner in which themes represented social contextual-level areas that differentiated the two groups can be found in Study 2. With respect to peer contexts, overall *Non-Improvers* were more likely to be influenced by the program participation of their peers, such that they were more likely to cite increased engagement (including attendance) as a result of peer/friend involvement in either their tutoring group or in the program as a whole. In contrast, *Improvers* were less likely to be influenced by the involvement of familiar peers in the program and were more personally rather than socially driven to participate in the program. Another differentiating factor at the familial level was the nature of parental (and familial) support of academic achievement. Although both groups of participants had parents who were generally very supportive of their academic achievement, *Improvers*' parents were more likely to provide instrumental supports, expose their youth to academic role models, and engage in reinforcing behaviors that likely encouraged youth's development of intrinsic rather than extrinsic motivation.

Broader School, Community and Cultural Level Contexts: Contextual Themes

A thorough discussion of the thematic findings and the manner in which themes represented broader level contextual areas differentiating the two groups can be found in Study 2. *Improvers* were generally more likely to have positive perceptions of their schools, were more engaged with their schools (e.g., extracurricular involvement) as well as with their neighborhoods/community (e.g., volunteer work), and comparatively exhibited a more solidly developed sense of cultural identity.

Summary of Results

Results from the analysis of the methods matrix indicated the following general finding. Based on the ten cases selected in Study 2 which were further examined in the current study, participants in both groups were generally comparable on the variables from Study 1 that were found to be potentially differentiating of *Improvers* and *Non-Improvers* (i.e., pre-intervention levels of language comprehension and motivation to read). Consideration of thematic findings from Study 2 indicated that multiple additional factors at different contextual levels represented supplemental sources of further differentiation of the groups. These factors extended beyond the individual demographic, cognitive, linguistic and behavioral factors identified at the quantitative research phase in Study 1.

Qualitative findings can be instrumental in elucidating broader contextual issues related to interventions (Weine et al., 2005; Creswell & Plano Clark, 2011). The integration of qualitative research methods with primarily quantitative methods allowed for a more comprehensive understanding of differentiating factors that could not have been captured in either research phase alone, elucidating the kinds of considerations and adaptations necessary to maximize student benefit from interventions such as the VLP, particularly in a multicultural context. What follows is a thorough discussion of the findings of this integrative research phase.

4.4 Discussion

This multi-phase mixed methods study was conducted to better understand factors that are potentially associated with differentiated reading intervention effects. Quantitative analysis in phase 1 (Study 1) addressed the effectiveness of the intervention relative to a control condition and factors predictive of change following the intervention were also examined. In order to better understand the multi-level, ecological contextual factors beyond the individual level factors examined in Study 1 that differentiate youth whose outcomes improved at posttest with those whose outcomes remained unchanged, qualitative analysis was conducted in phase 2 (Study 2) to gain additional insights using an ecological theoretical framework (Bronfenbrenner, 1979/2005). The current study represented an integration of findings from both prior research phases to generate a more comprehensive profile of individual and contextual factors that may be implicated in differential response to intervention, in particular the VLP reading intervention.

4.4.1 Comparison of findings across Studies 1 and 2

An integral component of the process of data integration was comparing findings across the studies.

Integrating data on differential effectiveness of the VLP. Subgroup analyses in Study 1 demonstrated that the VLP intervention was differentially effective. That is, treatment effects were significant and moderate in size for a subgroup of participants who displayed good

motivation to read prior to completing the intervention. Study 2 findings also confirmed that the intervention was differentially effective given the method used to differentiate outcome case clusters (i.e., *Improvers* vs *Non-Improvers*). Specifically, in Study 2 the reliable change index (RCI; Jacobson & Traux, 1991) was used to confirm that some participants indeed made statistically significant and clinically reliable improvement on reading comprehension (i.e., *Improvers*) while others did not exhibit significant outcome gains (i.e., *Non*-Improvers) as determined based on their relative comparison to a normative group. The Reliable Change Index (RCI) has been shown to be effective as an alternate measure of identifying intervention effects, although typically applied in the absence of a control group. Given that Study 1 indicated that overall intervention effects were nonsignificant relative to a comparison group, the RCI method, which is more sensitive to detecting both statistically and clinically significant change was used to identify intervention responders and non-responders. The fact that there were a considerable number of positive and negative cases of improvement identified across the entire sample, while sources of measurement error (i.e., practice effects) were simultaneously accounted for through the use of this methodology, was an indication that the intervention in fact resulted in differential effects.

Integrating data on predictive (and differentiating factors). Both studies attempted to extend beyond a focus on whether or not the intervention worked, to further exploration of whom it worked for and why. Study 1 identified that baseline motivation for reading and baseline language comprehension but not baseline word reading, were factors predictive of differential gains in reading comprehension outcomes after accounting for age, nonverbal ability and length of residence in Canada. Subgroup analyses indicated moderate size intervention effects for participants who reported good motivation to read pre-intervention. Study 2 findings pointed to superordinate themes of personal agency and motivation, in additional to various subthemes at

multiple contextual levels that were potentially associated with differential response to intervention. Results from data integration indicated that based on the ten cases selected in Study 2 and further employed for integrative analyses in the current study, participants in both groups were in fact comparable on the variables from Study 1 that were found to be potentially differentiating of *Improvers* and *Non-Improvers*. That is, the significant differentiating factors identified in Study 1 did not appear to differentiate the five Non-Improvers from the Improvers in this study. The fact this small selected subgroup of ten students held comparative quantitative profiles yet experienced differing outcomes supported the hypothesis that information from the first study alone would not comprehensively account for other potentially relevant differentiating factors. The comparability of these student's profiles made it imperative to look to the qualitative findings. The qualitative findings suggested that in addition to the individual factors identified in Study 1 (i.e., baseline reading comprehension and motivation to read), consideration of the following contextual factors may serve to further differentiate responders from non-responders, particularly when participants have similar cognitive-linguistic profiles: (1) at the individual level, achievement orientation, academic self-concept and sense of future aspirations; (2) at the program level, investment in positive program outcomes, tutoring group climate and participant resourcefulness; (3) at the social contextual level, nature of parental support and extent of reliance on peers; (4) at the broader school, community and cultural level, school perceptions, degree of engagement with school and community and sense of cultural identity. In addition to these, quantitative data also offered unique information about the potential role of parental education level in differentiating *Improvers* from *Non-Improvers*. Specifically, it indicated that *Improvers*' parents completed comparatively higher levels of formal education than the parents of Non-Improvers. Prior research indicates that parental education is associated with more instrumental parental academic support and overall academic achievement in children and youth

(Roundfield et al., 2018). Consistent with these findings, it is possible that higher levels of formal education among the parents of *Improvers* in this study contributed to their increased likelihood to provide instrumental support, to directly serve as academic role models for these students (or expose them to other familial models of academic success), as well as to support behaviors in their youth that were more consistent with intrinsic academic motivation.

Overall, these findings indicate that information from Study 2 supplemented conclusions from Study 1 to inform a more comprehensive picture of factors that may be implicated in differential response to the VLP reading intervention. The fact that the qualitative findings provided a deeper, enriched and contextualized narrative that enhanced the findings from phase one, highlighted the utility of focusing not only on whether the intervention worked but also for whom it worked and why.

4.4.2 General Discussion

If the goal of data integration in this study was to achieve triangulation of findings, it could be argued that the findings from both studies do not sufficiently converge. Within an expectation of convergence, the findings from both studies would need to be compared as follows.

Study 1 indicated that those participants who were more motivated to read represented the subgroup that made significant gains in vocabulary. Since vocabulary is highly related to reading comprehension, this subgroup would comprise the same youth expected to make better reading comprehension gains post intervention or at least over time (if outcomes were assessed longitudinally). Therefore if the primary researcher was seeking convergent findings, it would be reasonable to expect that those participants categorized as *Improvers* by virtue of study 2 criteria were more likely to be the same participants who reported higher motivation to read in Study 1. However, findings across both studies did not converge in this manner. That is, participants' reported motivation to read in Study 1 did not differentiate those categorized as *Improvers* from *Non-Improvers* in Study 2, as youth from both groups reported "good" motivation to read. In light of additional thematic findings from Study 2 suggesting that a major emerging differentiating factor between the groups was their comparative levels of underlying academic motivation in general, it could be interpreted that results from both studies are contradictory and perhaps warrant further investigation. Alternatively, it could equally be interpreted as evidence of convergence given that Study 1 findings indicated that well motivated readers did not in fact show significant intervention effects (relative to a control group) on reading comprehension outcomes *specifically* but instead on a vocabulary outcome.

Furthermore, even when findings are compared for convergence based solely on an identical reading comprehension outcome measure across both studies, findings yet again do not converge. Specifically, baseline language comprehension and motivation to read were shown in Study 1 to be associated with reading comprehension gains on the WRMT-R task. Therefore on the expectation of convergence, it would be expected that youth in the *Improvers* category would include a higher proportion of "good" language comprehenders (i.e., have higher scores than the sample mean on language comprehension), especially given that reading comprehension was the outcome determining case grouping in Study 2. However, similar to baseline motivation, baseline language comprehension did not differentiate this particular subgroup of ten participants.

In order to make sense of these seemingly conflicting findings, it is critical to recognize that convergence was not a goal of mixed method integration in this study, nor was convergence even possible given the multiple case study design employed in phase two. That is, in order to truly compare the studies with a goal of convergence, it would have been essential to achieve saturation in qualitative data analyses in Study 2 (to ensure maximal representativeness of the qualitative sample). More important, however, is the fact that the two separate study strands explored quite distinct factors that potentially differentiated responses to intervention. The purpose was for findings from both studies to supplement one another to form a broader understanding of the ecological profile of differentiating factors. This is a particularly important distinction to make given that the second study was based on a very small subsample of cases that cannot be purported to be representative of every student in the overall sample, nor should representativeness be expected given the heterogeneity of this culturally and linguistically diverse sample. Qualitative findings in Study 2 were not in any way based on attempting to achieve saturation of findings across participants. In fact one of the main strengths of this study's mixed method design was that it allowed for an in-depth examination of differentiating factors among small set of cases that only then ultimately resulted in the elucidation of the diversity (and at times seemingly unexpected similarity) among culturally and linguistically diverse immigrant and language minority youth. For example, it is possible (as indicated by the present study results) that if one isolates participants pre-intervention who appear to have similar baseline motivation and language comprehension profiles, they may not necessarily be expected to share similar outcomes post-intervention (as Study 1 results alone would suggest), but may in fact diverge significantly in outcomes depending on their individual broader environmental contexts.

Regardless of the fact that both studies addressed distinct yet complementary factors and essentially examined differing contexts, there was a unifying "meta-theme" of participant motivation identified. According to O'Cathain et al., (2010), a meta-theme is a theme that cuts across the qualitative and quantitative studies in a mixed method study. While Study 1 showed that *reading* motivation level at pretest made a difference in intervention effects, themes from Study 2 elaborated the ways in which multiple manifestations of motivation operated across all

relevant contextual areas to differentiate *Improvers* from *Non-Improvers*. Motivation is in itself a complex and multi-componential construct (Muho & Kurani, 2013). Further study into its role in differentiating intervention effects among adolescents who are culturally and linguistically diverse as well as socio-economically disadvantaged is highly warranted. Put together, these findings again highlight the importance of considering all possible contexts and adapting interventions based on the knowledge gained of the multiple factors that may be at play in differentiating participant outcomes.

If the entire sample of participants in the overall study were examined with their profiles of scores on Study 1 and Study 2 variables individually plotted against their outcomes, it is entirely possible that significantly predictive variables from Study 1 would characterize the individuals categorized as either *Improvers* or *Non-Improvers*. For example, those youth who were less motivated to read would overwhelmingly represent the same participants who did not improve post-intervention, with the opposite being the case for Improvers. However, the reality was that a small subset of ten participants were closely examined through multiple case study analysis, giving rise to the opportunity to closely examine particular possibilities of participant risk and resilience profiles that may have not have otherwise been considered.

Overall results from this study that explored multiple contexts navigated by culturally and linguistically diverse youth, suggest that the VLP reading intervention is differentially effective based on multiple factors at various contextual levels. These potential moderators of treatment effects that serve to promote risk or resilience were: 1) at the person-level, *pretest language comprehension skills, pretest motivation to read, achievement orientation, academic self-concept and sense of future aspirations*; 2) at the intervention program level, *investment in positive program outcomes, tutoring group climate and participant resourcefulness*; 3) at the peer and family social context, *reliance on peers and nature of parental support*, and 4) at the broader

school, community and cultural context, school perceptions, school and community engagement, and sense of cultural identity.

It is necessary to consider broader factors that may influence the likelihood of positive intervention effects, so that interventions can better accommodate for such factors. In fact, failure to account for multiple possibilities of factors influencing outcomes represents a disservice to youth, particularly those who are already academically at-risk based on socio-economic and related disadvantages. Every student stands the opportunity to benefit when interventions are optimized in ways that account for multiple possibilities and contextual sources of outcome differentiation. The findings from this study therefore highlight an opportunity for interventions to consider embracing the diversity of immigrant and language minority youth through use of innovative program designs.

This is especially important to employ for adolescent populations in light of recent research by Yeager, Dahl and Dweck (2017) suggesting that interventions often fail when traditionally less considered developmentally related factors such as the adolescent desire to be respected or accorded status, are not taken into account. Results from the current study help extend this finding to argue that considerations of adolescents' experiences and interactions with their broader contexts, in combination with individual, demographic and developmental factors collectively may contribute to optimal intervention benefits. In fact, taking such factors into account may indirectly serve to heighten youth motivation and engagement with academic intervention programs, factors that based on prior research often moderate key reading outcomes (e.g., Guthrie & Wigfield, 2000; Fredericks et al., 2004).

The inherent capacity for academic achievement of these academically at-risk youth cannot and should not be underestimated, particularly given their higher likelihood of resilience in the Canadian context, where immigration policies have resulted in a highly educated immigrant population (Statistics Canada, 2016a; OECD, 2018b). As evidenced by the parents of participants closely examined in this study, immigrant parents tend to place a high value on education, which is a protective factor for their youth. Recognition of these youth's potential allows for presenting them with appropriate challenges through intervention to exceed beyond environmentally-imposed limitations. The results of the evaluation of the VLP reading intervention for disadvantaged immigrant and language minority youth indicated that the intervention was not equally effective for all participants; instead, it was effective only for a select subgroup of participants exhibiting certain individual and broader contextual characteristics described in this study.

This author contends that there are some important lessons to be had in the seeming "failure" of this intervention to achieve uniform outcomes for the majority of participants. First, it allowed for an in-depth consideration of potential sources of measurement error that may have attenuated the likelihood of detecting intervention effects, as discussed in Study 1. Secondly, it reinforced the benefit of the mixed method approach employed whereby the research findings extended beyond exploration of intervention effects to examine factors related to differentiation of effects. Moreover, supplementation with qualitative methods allowed for a close examination and discovery of contextual areas potentially amenable to intervention related change that require additional study, or that at the very least that are worthy of consideration when aiming to tailor interventions to maximize student benefit. Findings from this study also highlight the potential loss of knowledge that may occur by passively accepting that intervention effects of any size are indication that interventions are serving urban society's increasing heterogeneous populations of students in uniform ways.

Limitations, Further Implications and Future Directions

Findings from this study are to an extent limited by the small sample size employed for the purpose of data integration. Thus although findings cannot be generalized to the broader population (particularly with respect to specific themes and meta-themes), this study offers critical insights about opportunities that may be considered in other intervention designs for maximizing intervention benefits in diverse populations. Overall, in helping to identify those at risk for poorer outcomes -- whether this is achieved by examining predictive models and differential effects as in Study 1, or by exploring differential contextual factors in Study 2 and then integrating the findings via mixed methodology – studies such as the present one can support the development of more effective interventions. This study also represents an innovative methodological approach for evaluating such programs, and would strongly benefit from replication.

References

- Anisef, P., Brown, R.S., Phythian, K., Sweet, R., & Walters, D. (2010). Early school leaving among immigrants in Toronto secondary schools. *Canadian Review of Sociology*, 47(2), 103-128.
- Archambault, I., Janosz, M., Dupere, V., Brault, M., & Mc Andrew, M. (2017). Individual, social, and family factors associated with high school dropout among low-SES youth: Differential effects as a function of immigrant status. *British Journal of Educational Psychology*. 87, 456-477.
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The Critical Role of Vocabulary Development for English Language Learners. *Learning Disabilities Research & Practice*, 20(1), 50 - 57.
- August, D., & Shanahan, T. (2006). Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth. Mahwah, NJ: Lawrence Erlbaum.
- August, D., & Shanahan, T. (Eds.). (2008). Developing reading and writing in second-language learners: Lessons from the Report of the National Literacy Panel on Language-Minority Children and Youth. New York: Routledge.
- August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness*, 2(4), 345–376.
- August, D., & Shanahan, T. (2010). Response to a Review and Update on Developing Literacy in Second-language Learners: Report of the National Literacy Panel on Language Minority Children and Youth. *Journal of Literacy Research*, 42(3), 341–348.
- Azimi, M., Iwenofu, L., Fraser, C. & Geva, E. (2013). Addressing social justice issues through literacy intervention in a vulnerable urban community [PowerPoint slides]. Social Justice Conference, Department of Applied Psychology and Human Development of the Ontario Institute for Studies in Education, Toronto, May 2013.
- Bagwell, C. L., & Schmidt, M. E. (2011). *Guilford series on social and emotional development*. *Friendships in childhood and adolescence*. New York, NY, US: Guilford Press.
- Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W.H. Freeman.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-Efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, 72, 187–206.
- Baxter, P. and Jack, S. (2008) Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13, 544-559.

- Becker, H., Roberts, G., & Voelmeck, W. (2003). Explanations for Improvement in Both Experimental and Control Groups. Western Journal of Nursing Research, 25(6), 746– 755.
- Bialystok, E., & Feng, X. (2011). Language proficiency and its implications for monolingual and bilingual children. *Language and Literacy Development in Bilingual Settings*, 121-138.
- Biemiller, A. (1999). Language and reading success. Cambridge, MA: Brookline Books.
- Biemiller, A. & Slonim, N. (2001) Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, 93, 498-520.
- Biemiller, A. (2005). Size and Sequence in Vocabulary Development: Implications for Choosing Words for Primary Grade Vocabulary Instruction. In E. Hiebert and M. Kamil (Eds.) *Bringing Scientific Research to Practice:Vocabulary*, Mahwah, NJ: Erlbaum.
- Biemiller, A., & Boote, C. (2006). An Effective Method for Building Meaning Vocabulary in Primary Grades. *Journal of Educational Psychology*, 98(1), 44-62.
- Biemiller, A. (2009). Words worth teaching: Closing the vocabulary gap. SRA/McGraw-Hill: Columbus, OH.
- Biemiller, A. (2012). Words for English-Language Learners. *TESL Canada Journal*, 29(6), 198-203.
- Bong, M., & Skaalvik, E. M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review*, 15, 1-40.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77-101.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist.* 32(7), 513-531.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (2005). Ecological systems theory (1992). In U. Bronfenbrenner (Ed.), Making human beings human: Biological perspectives on human development (pp. 106-173). Thousand Oaks, CA: Sage.
- Brown, B. B. (1990). Peer groups and peer cultures. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 171-196). Cambridge, MA, US: Harvard University Press.

- Buly, M. R., & Valencia, S. W. (2002). Below the Bar: Profiles of Students who Fail State Reading Assessments. *Educational Evaluation and Policy Analysis*, 24(3), 219–239.
- Cain, K. E., Bryant, P. E., & Oakhill, J. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal* of Educational Psychology, 96(1), 31-42.
- Cain, K. (2007). Syntactic awareness and reading ability: is there any evidence for a special relationship? *Applied Psycholinguistics*, 28, 679-694.
- Campbell, R., & Ahrens, C. (1998). Innovative Community Services for Rape Victims: An Application of Multiple Case Study Methodology. *American Journal of Community Psychology*, 26, 537-71.
- Carlisle J.F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Journal of Reading and Writing*, 12(3):169–190.
- Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, D. N., Lippman, D. N., ... White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of Englishlanguage learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, 39, 188–215.
- Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25-40). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Chall, J. S. (1996). Stages of reading development (2nd ed.). Fort Worth, TX: Harcourt Brace College Publishers.
- Chmiliar, L. (2010). Multiple-case designs. In A. J. MillsG. Durepos & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp. 583-584). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412957397.n216
- Corson, David. (1997). The Learning and Use of Academic English Words. *Language Learning*, 47, 671 718.
- Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(4), 213-238.
- Creswell, J.W. (1998). *Qualitative inquiry and research design Choosing among five traditions*. Thousand Oaks: SAGE Publications.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks: Sage Publication.
- Creswell, J.W. and Plano Clark, V.L. (2011) *Designing and conducting mixed methods research*. Sage, Thousand Oaks, California.

- Cronbach, L. & Snow, R. (1977). Aptitudes and Instructional Methods: A Handbook for Research on Interactions. New York: Irvington.
- Cruz de Quiros, A. M., Lara-Alecio, R., Tong, F., & Irby, B. J. (2010). The effect of a structured story reading intervention, story retelling and higher order thinking for English language and literacy acquisition. *Journal of Research in Reading*, 35(1), 87–113.
- Citizens for Public Justice. (2012). Poverty Trends Scorecard: Canada 2012. Retrieved from http://www.cpj.ca/files/docs/poverty-trends-scorecard.pdf
- Connor, C. M. (2011). Child characteristics-instruction interactions: Implications for students' literacy skills development in the early grades. In S. B. Neuman & D.K. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol. 3, pp 256–275). New York, NY: Guilford.
- Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cumming, A. (2012). Adolescent literacies in a multicultural context. New York: Routledge.

Cumming, A., & Geva, E. (2012a). Summary and recommendations. In A. Cumming (Ed.), *Adolescent literacies in a multicultural context* (pp. 133-148). New York: Routledge.

- Cumming, A., & Geva, E. (2012b). Purpose and approach. In A. Cumming (Ed.), *Adolescent literacies in a multicultural context* (pp. 1-22). New York: Routledge.
- Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge, England: Cambridge University Press.
- Cummins, J. (2012). The intersection of cognitive and sociocultural factors in the development of reading comprehension among immigrant students. *Reading and Writing*, 25(8), 1-18.
- Cummins, J. & Persad, R. (2014). Teaching through a multilingual lens: The evolution of EAL policy and practice in Canada. *Education Matters*, 2(1), 3-40.
- Cunningham, J. W., & Moore, D. W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Reading Behavior*, 25, 171–180.
- Dalecki, M., & Willits, F.K. (1991) Examining change using regression analysis: Three approaches compared. *Sociological Spectrum*, 11(2), 127-145.

Dalton, B., Proctor, C.P., Uccelli, P., Mo, E., & Snow, C. E. (2011). Designing for Diversity:

The Role of Reading Strategies and Interactive Vocabulary in a Digital Reading Environment for Fifth-Grade Monolingual English and Bilingual Students, *Journal of Literacy Research*, 43(1), 68–100.

- DeGenarro, D. (2018). The relationship between metacognitive strategies and reading comprehension in fifth grade students (Doctoral dissertation). ProQuest Dissertations and Theses database.
- Dei, G. (2008). Schooling as community: Race, schooling and the education of African youth. *Journal of Black Studies*, 38(3): 346-366.
- Dennis, J. M., Phinney, J. S., & Chuateco, L. I. (2005). The Role of Motivation, Parental Support, and Peer Support in the Academic Success of Ethnic Minority First-Generation College Students. *Journal of College Student Development*, 46(3), 223-236. http://dx.doi.org/10.1353/csd.2005.0023
- De Witte, K., Rogge N. (2013). Dropout from secondary education: All's well that begins well. *European Journal of Education*, 48(1), 131–149.
- Droop, M. & Verhoeven, L. (2003). Language proficiency and reading ability in firstand second-language learners. *Reading Research Quarterly*, 38(1), 78 - 103.
- Dunn, L.M., & Dunn, D.M. (2007). The Peabody Picture Vocabulary Test, Fourth Edition.
 Bloomington, MN: NCS Pearson, Inc.
 Dweck, C. S. (1986). Motivational processes affecting learning. American Psychologist, 41,1040–1048.
- Ebert, K. D., Kohnert, K., Pham, G., Rentmeester Disher, J., & Payesteh, B. (2014). Three treatments for bilingual children with primary language development: Examining crosslinguistic and cross-domain effects. *Journal of Speech, Language, and Hearing Research*, 57, 172–186.
- Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44, 414–448.
- Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: a mediational analysis. *Journal of Educational Psychology*, *91*, 549–563.
- Farnia, F., & Geva, E. (2011). Cognitive correlates of vocabulary growth in English language learners. *Applied Psycholinguistics*, 32(4), 711-738.
- Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, 36, 389-421.

Fernando, S. (2011). Pathways to Education Canada. International Symposium on Language and

Literacy in Immigrant Students [PowerPoint slides]. Ontario Institute for Studies in Education, Toronto, Ontario.

Field, A. (2009) Discovering Statistics Using SPSS. 3rd Edition, Sage Publications Ltd., London.

- Fien, H., Anderson, D., Nelson, N.J., Kennedy, P., Baker, S. K., & Stoolmiller, M. (2018). Examining the impact and school-level predictors of impact variability of an 8th grade reading intervention on at-risk students' reading achievement. *Learning Disabilities Research and Practice*, 33(1), 37-50.
- Fletcher, J., & Vaughn, S. (2009). Response to Intervention: Preventing and Remediating Academic Difficulties. *Child Development Perspectives*, 3, 30-37. https://doi.org/10.1111/j.1750-8606.2008.00072.x
- Follmer, J. & Sperling, D. (2018). Interactions between reader and text: Contributions of cognitive processes, strategy use, and text cohesion to comprehension of expository science text. *Learning and Individual Differences*, 67, 177-187.
- Fredricks, J., Blumenfeld, P., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109.
- Fuchs, L.S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research and Practice*, 18(3), 172-186.
- Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What Is Intensive Instruction and Why Is It Important? *Teaching Exceptional Children*, 46(4), 13-18.
- Fuchs, D., Hendricks, E., Walsh, M., Fuchs, L., Gilbert, J., Patton, S., ...& Peng, P. (2017). Evaluating a Multidimensional Reading Comprehension Program and Reconsidering the Lowly Reputation of Tests of Near-Transfer. *Learning Disabilities Research and Practice*. 33(1), 11-23.
- Gambrell, L.B., Malloy, J.A., & Mazzoni, S.A. (2011). Evidence based best practices in comprehensive literacy instruction. In L.M. Morrow & L.B. Gambrell (Eds.), *Best* practices in literacy instruction (4th ed., pp. 11–36). New York: Guilford.
- Geva, E., & Ryan, E. B. (1993). Linguistic and memory correlates of academic skills in first and second languages. *Language Learning*, 43, 5-42.
- Geva, E. (2006). Second-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: A report of the National Literacy Panel on Language-Minority Children and Youth* (pp.123-139). Mahwah, NJ: Erlbaum.
- Geva, E., Farnia, F., Chen, X., & Gottardo, A. (2011). Altering the academic achievement trajectory of recent adolescent immigrants: An after school community-based intervention targeting language and reading related strategic knowledge (Proposal).

(Available from the Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, ON, M5S 1V6).

- Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing: An Interdisciplinary Journal*, 25(8), 1819-1845.
- Geva, E. & Wiener, J. (2015). *Psychological Assessment and Intervention with Culturally and Linguistically Diverse Children and Adolescents*. New York: Springer.
- Gibson, M., & Ogbu, J. (1991). *Minority Status and Schooling: A Comparative Study of Immigrant and Involuntary Minorities* (New York, London).
- Goodenow, C., & Grady, K. E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *Journal of Experimental Education*, 62(1), 60-71.http://dx.doi.org/10.1080/00220973.1993.9943831
- Goodrick, D. (2014). Comparative Case Studies: Methodological Briefs Impact Evaluation No.
 9. Florence, Italy: United Nations Children's Fund (UNICEF). Retrieved from https://www.unicef-irc.org/publications/pdf/brief_9_comparativecasestudies_eng.pdf
- Gottardo, A., & Mueller, J. (2009). Are first and second language factors related in predicting school language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, 101, 330-334.
- Gough, P.B., & Tunmer. W.E. (1986). Decoding, reading, and reading disability. *RASE: Remidial & Special Education*, 7, 6-10.
- Green, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting High School Students' Cognitive Engagement and Achievement: Contributions of Classroom Perceptions and Motivation. *Contemporary Educational Psychology*, 29, 462-482. http://dx.doi.org/10.1016/j.cedpsych.2004.01.006
- Green, S. B., & Salkind, N. J. (2011). Using SPSS for Windows and Macintosh: Analyzing and understanding data (Sixth). New Jersey: Pearson Education, Inc.
- Green, J., Liem, G., Martin, A., Colmar, S., Marsh, H., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of adolescence*, 35, 1111-1122.
- Guruge, S., & Collins, E. (2008). Working with immigrant women: Issues and strategies for mental health professionals. Toronto: Centre for Addiction and Mental Health.
- Guthrie, J., & Wigfield, A. (2000). Engagement and motivation in reading. In M. Kamil, P. Mosenthal, P. Pearson & R. Barr (Eds.), *Handbook of reading research* (3rd ed.,). Mahwah, NJ: Lawrence Earlbaum Associates.

- Guthrie, J. T. (2004). Teaching for Literacy Engagement. *Journal of Literacy Research*, 36(1), 1–30. https://doi.org/10.1207/s15548430jlr3601_2
- Guthrie, J., Mcrae, A., & Lutz Klauda, S. (2007). Contributions of Concept-Oriented Reading Instruction to Knowledge About Interventions for Motivations in Reading. *Educational Psychologist*, 42, 237-250.
- Harris, E. D. (1988). Matrix Analogies Test Extended Form: a Review. Canadian Journal of School Psychology, 4(1), 77–79.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD, US: Paul H Brookes Publishing.
- Hoover, W. A., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal*, 2(2), 127-160. http://dx.doi.org/10.1007/BF00401799
- Hou, F. & Zhang, Qi. (2015). Regional differences in the educational outcomes of young immigrants. *Insights on Canadian Society*. November. Statistics Canada Catalogue no. 75-006-X.
- Human Resources Skills and Development Canada (HRSDC). (2008). *Reading skills of young immmigrants in Canada: the effects of duration of residency, home language exposure and schools.* Quebec, CA. Learning Research Series, Human Resources and Social Development Canada. Retrieved from http://www.yorku.ca/pathways/literature/Literacy/gluszynski%2008.pdf
- Hutchinson, J.M., Whiteley, H., Smith, C., & Connors, L. (2003). The early identification of dyslexia: children with English as an additional language. *Dyslexia*, 10(3):179-95.
- Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2015). Differential effects of a systematic vocabulary intervention on adolescent language minority students with varying levels of English proficiency. *International Journal of Bilingualism*, 19(3), 314–332.
- Ivey, G., & Broaddus, K. (2001). "Just plain reading": A survey of what makes students want to read in middle school classrooms. Reading Research Quarterly, 36(4), 350–377.
- Iwenofu, L. (2013). Examining student engagement in literacy intervention: Voices of adolescents living in an urban marginalized community (Masters thesis). Retrived from https://tspace.library.utoronto.ca/bitstream/1807/42626/1/Iwenofu_Linda_201311_MA_t hesis.pdf
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19.

Janosz, M., LeBlanc, M., Boulerice, B., & Tremblay, R. E. . (1997). Disentangling the weight of

school dropout predictors: A test on two longitudinal samples. *Journal of Youth and Adolescence*, 26,733–762.

- Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language children and its role in word recognition ability. *Applied Psycholinguistics*, 30(1), 153 185.
- Jeon, Eun Hee & Yamashita, Junko. (2014). L2 Reading Comprehension and Its Correlates: A Meta-Analysis. Language Learning. 64(1), 160-212.
- Johnston, P.H. (2004). *Words: How our language affects children's learning*. Portland, ME: Stenhouse.
- Jung, T., & Wickrama, K.A.S. (2008). An introduction to latent class growth analysis and growth mixture modelling. Social and Personality Psychology Compass, 2(1), 302-317.
- Juvonen, J. (2006). Sense of Belonging, Social Bonds, and School Functioning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 655-674). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Kamil, M., Mosenthal, P., Pearson, P. D., & Barr, R., (Eds.). (2000). Handbook of Reading Research, Volume III. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Kendeou, P., Savage, R., & van den Broek, P. (2010). Revisiting the simple view of reading. *British Journal of Educational Psychology*, 79(2), 353-370.
- Kieffer, M. J., & Lesaux, N. K. (2008). The Role of Derivational Morphology in the Reading Comprehension of Spanish-Speaking English Language Learners. *Reading and Writing*, 21, 783-804.
- Kitto, S., Chesters, J., & Grbich, C. (2008). Quality in qualitative research. *The Medical Journal* of Australia, 188(4), 243–246.
- Klauda, S.L. & Guthrie, J.T. (2015) Comparing relations of motivation, engagement, and achievement among struggling and advanced adolescent readers. *Reading and Writing*, 28(2), 239-269.
- Kratochwill, T. R., & Shernoff, E. S. (2003). Evidence-Based Practice: Promoting Evidence-Based Interventions in School Psychology. *School Psychology Quarterly*, 18(4), 389-408. http://dx.doi.org/10.1521/scpq.18.4.389.27000
- Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, 55(1), 170-183. http://dx.doi.org/10.1037/0003-066X.55.1.170
- Lervåg, Arne & Grøver Aukrust, Vibeke. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension between first and second language learners. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51, 612-20.

- Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal, 19*(1), 99-131.
- Lesaux, N., Geva, E., Koda, K., Siegel, L., & Shanahan, T. (2008) Development of literacy in second language learners. In August & Shanahan (eds.) (2008). *Developing reading and writing in second-language learners: Lessons from the Report of the National Literacy Panel on Language-Minority Children and Youth.* New York: Routledge.pp.27-59.
- Lesaux, N.K., & Kieffer, M.J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, 47(3), 596–632.
- Lesaux, N.K., Kieffer, M.J., Faller, S.E., & Kelley, J.G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196-228.
- Lesaux, N., Harris, J., & Sloane, P. (2012). Adolescents' motivation in the context of an academic vocabulary intervention in urban middle school classrooms. *Journal of Adolescent & Adult Literacy*, 56(3), 231-240.
- Lesaux, N.K., Kieffer, M.J., Kelley, J.G., & Harris, J.R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, 51(6), 1159-1194.
- Li, M., & Kirby, J. (2014). Unexpected Poor Comprehenders Among Adolescent ESL Students. *Scientific Studies of Reading*, 18, 75-93.
- Lin, D., Ramirez, G., Wilson, J. & Geva, E. (2012). Bridging lexical knowledge and literacy. In A. Cumming (Ed.), Adolescent literacies in a multicultural context (pp. 102-117). New York: Routledge.
- Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage.
- Linley, R. (2004). What school factors contribute to underachievement? The case of caribbean students in Toronto secondary schools (Unpublished master's thesis). University of Ottawa. Ottawa, ON.
- Lin-Siegler, X., Dweck, C. S., & Cohen, G. L. (2016). Instructional interventions that motivate classroom learning. *Journal of Educational Psychology*, 108(3), 295-299. http://dx.doi.org/10.1037/edu0000124
- Little, R.J.A. & Rubin, D.B. (2002). *Statistical analysis with missing data*. (Second Edition.) New York: Wiley.
- Lochner, L. & Moretti, E. (2004) The effect of education on crime: evidence from prison

inmates, arrests, and self-reports, American Economic Review, 94,155–189.

- Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences*, 21(1), 124-128.
- Lopez, G. (2001). The value of hard work: Lessons on parent involvement from an (im)migrant household. *Harvard Educational Review*, *71*, 416-437.
- Lovett, M. W., Palma, M. D., Frijters, J., Steinbach, K., Temple, M., Benson, N., & Lacerenza, L. (2008). Interventions for Reading Difficulties: A Comparison of Response to Intervention by ELL and EFL Struggling Readers. *Journal of Learning Disabilities*, 41(4), 333–352. https://doi.org/10.1177/0022219408317859
- Low, P., & Siegel, L.S. (2005) A comparison of the cognitive processes underlying reading comprehension in native English and ESL speakers. Written Language and Literacy, 8(2), 207-231.
- Lubliner, S., & Smetana, L. (2005). The Effects of Comprehensive Vocabulary Instruction on Title I Students' Metacognitive Word-Learning Skills and Reading Comprehension. *Journal of Literacy Research*, 37(2), 163–200.
- Luthar, S. S., & Cicchetti, D. (2000). The construct of resilience: Implications for interventions and social policies. *Development and Psychopathology*, 12(4), 857-885. https://doi.org/10.1017/S0954579400004156
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). Gates-MacGinitie reading tests: Manual for scoring and interpretation. Itasca, IL: Riverside.
- Mahoney, J. L., & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology*, 33, 241–253.
- Malcolmson, J. D. "What Works in Youth Literacy and Why?: A Literature Review and Discussion." Unpublished paper prepared for Youth Initiatives Directory, Human Resources and Development Canada, October 2001.
- Mancilla-Martinez, J. (2010). Word meanings matter: Cultivating English vocabulary knowledge in fifth-grade Spanish-speaking language minority learners. TESOL Quarterly, 44, 669– 699.
- Mancilla-Martinez, J., & Lesaux, N.K. (2010). Predictors of Reading Comprehension for Struggling Readers: The Case of Spanish-speaking Language Minority Learners. *Journal* of educational psychology, 102 3, 701-711.
- Marsh, H. W. (2007). Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology. Leicester, UK: British Psychological Society.
- Masten, A. S., & Garmezy, N. (1985). Risk, vulnerability, and protective factors in

developmental psychopathology. In B. B. Lahey, & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 8, pp. 1-52). New York, NY: Plenum Press.

- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, 56(3), 227-238.http://dx.doi.org/10.1037/0003-066X.56.3.227
- Masten, A. S., & Powell, J. L. (2003). A Resilience Framework for Research, Policy and Practice. In S. S. Luthar (Eds.), Resilience and Vulnerability: Adaptation in the Context of Childhood Adversities (pp. 1-26). New York: Cambridge University Press.
- Masten, A. S. (2004). Regulatory processes, risk, and resilience in adolescent development. In R.
 E. Dahl & L. P. Spear (Eds.), Annals of the New York Academy of Sciences: Vol. 1021.
 Adolescent brain development: Vulnerabilities and opportunities (pp. 310-319). New York, NY, US: New York Academy of Sciences.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive-achievement relations: What we ave learned from the past 20 years of research. *Psychology in the Schools*, 47, 651-675.
- McKool, S. (2007). Factors that influence the decision to read: An investigation of fifth grade students'out-of-school reading habits. *Reading Improvement*, 44(3), 111-131.
- Meece, J.L., Anderman, E.M., & Anderman, L.H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, 57, 487–503.
- Merriam, S.B. (1998). Qualitative research and case study application in education. San Francisco, CA: Jossey-Bass.
- Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformativeemancipatory perspective. In A.Tashakkori & C.Teddlie (Eds.), Handbook of mixed methods in social and behavioral research (pp. 135–164). Thousand Oaks, CA: Sage.
- Mertens, D. M. (2007). Transformative Paradigm: Mixed Methods and Social Justice. *Journal of Mixed Methods Research*, 1(3), 212–225. https://doi.org/10.1177/1558689807302811
- Miciak, J., Roberts, G., Taylor, W.P., Solis, M., Ahmed, Y., Vaughn, S., & Fletcher, J.M. (2018).

The effects of one vs. two years of intensive reading intervention implemented with late elementary struggling readers. *Learning Disabilities Research and Practice*, 33(1), 24-36.

- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage Publications.
- Mokhtari, K., & Reichard, C. (2002). Assessing Students' Metacognitive Awareness of Reading Skills. *Journal of Educational Psychology*, 94, 249-259.
- Muho, A., & Kurani, A. (2013). Components of motivation to learn from a psychological

perspective. Academic Journal of Interdisciplinary Studies, 2(4), 173-180.

- Naglieri, J. A. (1985). *Matrix analogies test—Expanded form*. San Antonio, TX: The Psychological Corporation.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91–108.
- Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, 20(7), 691-719.
- Nation, Paul & Kyongho, Hwang. (1995). Where would general service vocabulary stop and special purposes vocabulary begin?. *System*, 23, 35-41.
- National Reading Panel (NRP). (2000). *Teaching children to read : an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, D.C.: National Institute of Child Health and Human Development, National Institutes of Health.
- Nazarieh, M. (2015). A brief history of achievement goals orientations theory and its development. *BEST: Journal of Humanities, Arts, Medicine and Sciences*, 1(2), 49-58.
- Neuman, S.B., & Dickinson, D.K. (2003). *Handbook of early literacy research*. New York, NY: Guilford Press.
- O'Brien, L. M., Paratore, J. R., Leighton, C. M., Cassano, C. M., Krol-Sinclair, B., & Green, J. G. (2014). Examining Differential Effects of a Family Literacy Program on Language and Literacy Growth of English Language Learners With Varying Vocabularies. *Journal of Literacy Research*, 46(3), 383–415.
- O'Cathain A., Murphy, E., & Nicholl J. (2010). Three techniques for integrating data in mixed methods studies. *British Medical Journal*, 341:c4587.
- O'Connor, M., Geva, E., & Koh, P. (2018). Examining Reading Comprehension Profiles of Grade 5 Monolinguals and English Language Learners Through the Lexical Quality Hypothesis Lens. *Journal of Learning Disabilities*.
- Organization for Economic Co-operation and Development [OECD]. (2018a). OECD Economic Surveys of Canada: Key Insights. Retrieved from: www.oecd.org/eco/surveys/economic-survey-canada.htm *OECD Economic Surveys*, Canada.
- OECD. (2018b). The Resilience of Students with an Immigrant Background: Factors that Shape Well-being. DOI:https://dx.doi.org/10.1787/9789264292093-en
- Ogbu, J. (1991). Minority Coping Responses and School Experience. *The Journal of Psychohistory*, 18, 433-456.

- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554-566.
- Ouellette, G., & Beers, A. (2010) A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, 23, 189–208.
- Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Pasquarella, A., Gottardo, A., & Grant, A. (2012). Comparing factors related to reading comprehension in adolescents who speak English as a first (L1) or second (L2) language. *Scientific Studies of Reading*, 16(6), 475-503.
- Pasquarella, A., Fraser, C., Kornacki, T., Iwenofu, L., Azimi, M., Farnia, F., & Geva, E. (2013). Vocabulary Learning Project: Tutor Training Manual. (Available from the Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, ON, M5S).
- Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*. Thousand Oaks, CA: Sage Publications.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357–383.
- Portes, A., & MacLeod, D. (1996). Educational progress of children of immigrants: The roles of class, ethnicity, and school context. *Sociology of Education*, 69(4), 255-275. http://dx.doi.org/10.2307/2112714
- Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-Speaking Children

Reading in English: Toward a Model of Comprehension. *Journal of Educational Psychology*, *97*(2), 246-256.

Proctor, C.P., Dalton, D., Uccelli, P., Biancarosa, G., Mo, E., Snow, C.E., & Neugebauer, S.

(2011). Improving Comprehension Online (ICON): Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing: An Interdisciplinary Journal*, 24(5), 517-544.

- Ram, N., & Grimm, K. (2009). Growth Mixture Modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, 33(6), 565–576.Int J Behav Dev. 2009; 33(6): 565–576. https://doi.org/10.1177/0165025409343765
- Rossi, P., Freeman, H., & Lipsey, M. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage Publications.

- Roundfield, K. D., Sánchez, B., & McMahon, S. D. (2018). An Ecological Analysis of School Engagement Among Urban, Low-Income Latino Adolescents. *Youth & Society*, 50(7), 905–925. https://doi.org/10.1177/0044118X16639986
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons Inc., New York.
- Rubinstein-Ávila, E. (2003). Facing Reality: English Language Learners in Middle School Classes. *English Education*, 35(2), 122-136.
- Rumberger, R. W. (2011). *Dropping out: Why students drop out of high school and what can be done about it.* Cambridge, MA: Harvard University Press.
- Rutter, M. (2000). Resilience reconsidered: Conceptual considerations, empirical findings, and policy implications. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (pp. 651-682). New York, NY, US: Cambridge University Press.
- Rutter, M. (2006). Implications of Resilience Concepts for Scientific Understanding. Annuals of the New York Academy of Sciences, 1094, 1-12.
- Ryan, A. (2000). Peer groups as a context for the socialization of adolescents' motivation, engagement, and achievement in school. *Educational Psychologist*, 35(2), 101-111.
- Santoro, L. E., Jitendra, A. K., Starosta, K., & Sacks, G. (2006). Reading well with Read Well. *Remedial and Special Education*, 27(2), 105–115.
- Scarcella, Robin. (2003). *Academic English: A Conceptual Framework*. University of Minnesota Linguistic Minority Research Institute.
- Schafer, J.L. & Graham, J.W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*,7, 147–177.
- Schneider, B. H., Wiener, J., & Murphy, K. (1994). Children's friendships: The giant step beyond peer acceptance. *Journal of Social and Personal Relationships*, 11(3), 323-340. http://dx.doi.org/10.1177/0265407594113001
- Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching English-language learners. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 415–488). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Skinner, E. A., Kindermann, T. A., Connell, J. P., & Wellborn, J. G. (2009). Engagement and disaffection as organizational constructs in the dynamics of motivational development. In K. R. Wenzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (pp. 223-245). New York, NY, US: Routledge/Taylor & Francis Group.

- Snow, C.E., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Snow, C.E, & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York: Carnegie Corporation of New York.
- Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, 2(4), 325–344.
- Snyder, Elizabeth & Witmer, Sara & Schmitt, Heather. (2017). English language learners and reading instruction: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth*, 61(2), 136-145.
- Solga, H. (2002). Stigmatization by negative selection. Explaining less-educated people's decreasing employment opportunities. *European Sociological Review*, 18, 159–178.
- Spycher, P. (2009). Learning academic language through science in two linguistically diverse kindergarten classes. *The Elementary School Journal*,109, 359–379.
- Stake, R. (1995). The art of case study research. Thousand Oaks, CA: Sage Publications.
- Stake, R. E. (2005). Qualitative Case Studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 443-466). Thousand Oaks, CA, : Sage Publications Ltd.
- Stake, R. E. (2006). Multiple Case Study Analysis. New York: Guilford Press.
- Stanovich, Keith E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 22, 360-407.
- Stanovich, P. J., & Jordan, A. (1998). Canadian teachers' and principals' beliefs about inclusive education as predictors of effective teaching in heterogeneous classrooms. *Elementary School Journal*, 98, 221-238.
- Stahl, S.A., & Nagy, W.E. (2006). *Teaching word meanings*. Mahwah, NJ: Erlbaum.
- Statistics Canada. (2013). Immigration and Ethnocultural Diversity in Canada. Statistics Canada Catalogue no. 99-010-X. Ottawa Ontario. Analysis Series, 2011 National Household Survey. https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001eng.pdf
- Statistics Canada. (2016a). Immigration and ethnocultural diversity: Key results from the 2016 census. Statistics Canada Catalogue no. 11-001-X. Ottawa Ontario. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/171025/dq171025b-eng.pdf
- Statistics Canada. (2016b). Census of Population, 2016: Ottawa, Ontario. Regent Park, 2016 Neighborhood Profile. Retrieved from https://www.toronto.ca/ext/sdfa/Neighbourhood%20Profiles/pdf/2016/pdf1/cpa72.pdf

- Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement: An ecological perspective. *American Psychologist*, 47(6), 723-729. http://dx.doi.org/10.1037/0003-066X.47.6.723
- Storch, S. A., & Whitehurst, G. J. (2002). Oral Language and Code-Related Precursors to Reading: Evidence from a Longitudinal Structural Model. *Developmental Psychology*, 38, 934-947.
- Stutz, F., Schaffner, E., & Schiefele, U. (2016). Relations among reading motivation, reading amount, and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 45, 101-113.
- Tarar, Meisinger & Dickens (2015). Test Review: Test of Word Reading Efficiency–Second Edition (TOWRE-2) by Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. Canadian *Journal of School Psychology*, 30 (4): 320-326.
- Therrien, W. & Cook, B. (2018). Introduction to special issue: Null effects and publication bias in learning disabilities research. *Learning Disabilities Research and Practice*, 33 (1), 5-10.
- Tilleczek, K., & Campbell, V. (2013). Barriers to youth literacy: Sociological and Canadian insights. *Language and Literacy*, 15(2), 77-100.
- Tilleczek, K. (2008). Building Bridges: Transitions from Elementary to Secondary School. *Education Canada*, 48 (1), 68-71.
- Tong, F., Irby, B. J., Lara-Alecio, R., & Koch, J. (2014). Integrating literacy and science for English language learners: From learning-to-read to reading-to-learn. The Journal of Educational Research, 107(5), 410–426.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). Test of Word Reading Efficiency– Second Edition (TOWRE-2). Austin, TX: Pro-Ed.
- Toronto Community Housing (2007). *Regent Park Social Development Plan*. Retrieved from https://www.toronto.ca/legdocs/mmis/2007/ex/bgrd/backgroundfile-7300.pdf
- Toronto District School Board. (2008). Toronto district school board urban diversity strategy: Focus on student achievement. (Staff Report). Toronto: Toronto District School Board (TDSB). Retrieved from: http://www.tdsb.on.ca/wwwdocuments/parents/special_education_2/docs/Urban%20Dive rsity %20Strategy.pdf
- Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing*, 22(9), 993–1019.

Unrau, N., & Schlackman, J. (2006). Motivation and its relationship with reading achievement in
an urban middle school. Journal of Educational Research, 100, 81-101.

- Vadasy, P. F., & Sanders, E. A. (2010). Efficacy of supplemental phonics based instruction for low-skilled kindergarteners in the context of language minority status and classroom phonics instruction. *Journal of Educational Psychology*, 102, 786–803.
- van Ginkel, J.R. & Kroonenberg, P. (2014). Analysis of variance of multiply imputed data. *Multivariate Behavioral Research*, 49, 78-91.
- Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., ... Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal*, 107(2), 153–180.
- Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research* on Educational Effectiveness, 2(4), 297–324.
- Vaughn, S., Klingner, J., Swanson, E., Boardman, A., Roberts, G., Mohammed, S., & Stillman-Spisak, S. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, 48(4), 1-27.
- Verhoeven, L. (2000). Components in Early Second Language Reading and Spelling. Scientific Studies of Reading, 4, 313-330.
- Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, 22, 407 423.
- Verhoeven, L., & van Leeuwe, J. (2011). The simple view of second language reading throughout the primary grades. *Reading and writing*, 25(8), 1805-1818.
- Wang, R., & Ware, J. (2013). Detecting moderator effects using subgroup analyses. Prevention Science, 14(2), 111-120.
- Weine, S., Knafl, K., Feetham S., Kulauzovic, Y., Klebic, A., Sclove, S., ...Spahovic, D. (2005). A mixed methods study of refugee families engaging in multiple family groups. *Family Relations*, 54(4), 558-568.
- Wentzel, K. R. (2005). Peer Relationships, Motivation, and Academic Performance at School. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 279-296). New York, NY, US: Guilford Publications.
- Werner, E. E., & Smith, R. S. (1992). Overcoming the Odds: High Risk Children from Birth to Adulthood. Ithaca, NY: Cornell University Press.
- Wiener, J., & Daniels, L. (2016). School Experiences of Adolescents With Attention Deficit/Hyperactivity Disorder. *Journal of Learning Disabilities*, 49(6), 567-581.

https://doi.org/10.1177/0022219415576973

- Wigfield, A. & Guthrie, J.T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89, 420-432.
- Wigfield, A., Eccles, J.S., & Rodriguez, D. (1998). The development of children's motivation in school contexts. *Review of Research in Education*, 23(1), 73–188.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.
- Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 15-31). San Diego, CA: Academic Press.
- Williams, K.T. (2007). Expressive Vocabulary Test (EVT-2). San Antonio, TX: Pearson.
- Wilson, J. S. (2012). Social support networks for literacy engagement among culturally diverse urban adolescents. (University of Toronto (Canada)). ProQuest Dissertations and Theses, 309. Retrieved from http://search.proquest.com/docview/1346237450?accountid=14771.
- Wilson, S.J., & Lipsey, M. (2014). Exploring variability in treatment effects [PowerPoint presentation].Retrieved from https://my.vanderbilt.edu/evte/files/2013/10/Wilson-SRA-2014.pdf
- Wise, E. (2004). Methods for analyzing psychotherapy outcomes: a review of clinical significance, reliable change, and recommendations for future directions. *Journal of* 155 *Personality Assessment*, 82(1), 50–59.
- Won Jun, S., & Watanabe, Y. (2012). *Reading practices and achievement*. In A. Cumming (Ed.), Adolescent literacies in a multicultural context (pp. 66-73). New York: Routledge.
- Woodcock, RN. (1987). Woodcock Reading Mastery Tests-revised examiner's manual. Circle Pines, MN: American Guidance Service.
- Yaghoub-Zadeh, Z., Farnia, F., & Geva, E. (2012). Toward Modeling Reading Comprehension and Reading Fluency in English Language Learners. *Reading and Writing: An Interdisciplinary Journal*, 25, 163-187.
- Yeager, D., Dahl, R., Dweck, C. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, 13 (1), 101-122.
- Yesil Dagli, Ummuhan. (2010). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly*, 26, 15-29.

Yin, R.K. (2003) Case Study Research: Design and Methods. 3rd Edition, Sage, Thousand Oaks. Zolli, A., & Healy, AM. (2012). Resilience: Why things bounce back. New York: Free Press.

Appendices

Appendix A

Comparison of Program Implementation Across Intervention Cohorts

Comparison of manual versions

During this first cohort, a pilot version of the intervention manual was used (version 1). During the second and third intervention cohorts, a slightly revised manual was implemented (version 2). The two manual versions were compared. The following table summarizes key aspects that were modified for version 2 of the manual.

Program Aspect	Differences between manual versions
Vocabulary cards	Version 2 has clearer instructions indicated on
	the cards vs. version 1
	Vocabulary card content modified
Rewards program	Directions for the implementation of a new
	rewards program was introduced in version 2,
	whereby participants can accumulate points
	toward earning a reward at the end of the
	program
Session icebreakers	Version 1 indicated "suggested" icebreaker
	activities; Version 2 clearly articulated specific
	icebreakers to be implemented for each session
Logs of vocabulary words taught	Version 2 introduced pages for tutors to log a
	"running vocabulary list" of new words taught
	and learned by their group of students

For the purposes of this study, the intervention as implemented in the first cohort (version 1) was considered to represent a distinct, pilot version of implementation of the intervention.

Comparison of tutor training procedures

Tutors across all cohorts comprised adult volunteers who were trained in the effective implementation of the VLP intervention, in accordance with the manual version being used. Tutor training for all cohorts was conducted by senior graduate students, in consultation with and under the supervision of the primary study investigators. For Cohort 1, there were two graduate trainers, and training was completed in a single session. For Cohorts 2 and 3, there were four training for Cohort 1 also differed from Cohort 2 and 3 with respect to aspects of the program that were emphasized. For Cohort 1, the training format involved providing considerable background and theoretical information about reading instruction and intervention, with less focus on practical instruction on how to teach the reading strategies. Given consensus of the research team that this approach was likely ineffective, the training approach for Cohorts 2 and 3

was significantly modified such that its primary focus was on explicitly teaching tutors about the strategies. Complete lesson plans were reviewed so that tutors could grasp the program's structure, with opportunities for tutors to practice leading a mock program session with each other.

Comparison of program adherence rates

A detailed description of procedures for determining program adherence/fidelity was provided in Study 1 (section 2.2.7). In general, fidelity assessments involved rating tutors' ability to complete core program aspects outlined in the manual, per session checked. Overall adherence rates were as follows for cohorts 1, 2 and 3 respectively: 82%, 92% and 87%.

Evaluation of intervention effects from pre to post on outcome variables (with Cohort 1 data included)

Outcome analyses for treatment effects were conducted with data from intervention cohort 1 to explore comparability of findings had this cohort been included in overall analyses (see Table A1). Results indicated similar intervention effects (i.e., null effects) to analyses conducted with the entire sample of intervention participants (i.e., cohorts 2 and 3) across all relevant outcome measures.

Table A1

					Group
Outcome	Outcome			Posttest	effect/ANOVA
variable	measure	Group	Pretest M(SD)	M(SD)	Group x Time effect
Vocabulary	Expressive vocabulary <i>df</i> (1,230)	Intervention Control	123.63 (24.51) 124.43 (19.08)	128.55 (23.79) 130.72 (18.48)	F=1.17, p =0.28, η^2 =0.007
	Receptive vocabulary <i>df</i> (1,230)	Intervention Control	163.77 (32.40) 171.29 (19.70)	168.20 (31.27) 173.76 (20.58)	F=1.28, p =0.26, η^2 =0.006
	Academic vocabulary df(1,230)	Intervention Control	16.21 (3.59) 17.02 (2.29)	16.37 (3.04) 16.76 (2.49)	F=1.58, <i>p</i> =0.21, η ² =0.007
Reading Comprehension	WRMT-R Passage Comp. <i>df</i> (1,230)	Intervention	44.11 (9.70)	45.01 (9.04)	F=0.64, p =0.43, η^2 =0.004

Results of Mixed ANOVAs for Intervention Effects (Intervention Cohorts 1-3, n=127 vs. Controls, n=105)

	Control	43.58 (6.50)	45.29 (8.36)	
GMRT Reading Comp. <i>df</i> (1,230)	Intervention	23.59 (10.97)	24.40 (11.16)	F=0.26, p =0.61, η^2 =0.001
	Control	25.09 (9.29)	25.44 (9.10)	

Note. Effect sizes reported are partial eta squared.

Appendix B

Measures: Family Background Questionnaire, Pre-Interview Questionnaire and Semi-structured Student Interview Guide

B.1 Family Background Questionnaire

Family Background Questionnaire

We need to get from you some background information about your son/daughter who is participating in the project, in order to be able to better understand the factors that influence your child's ability to learn and how they feel about themselves. We would greatly appreciate if you provide us with some information about your child's learning history and development, your family composition, employment and educational background.

	Today's date: day / month / year
1. My child's name is	, born on day / month / year
Child's gender: Female O Male O	
2. Name of current school	
 3. In what grade is your child enrolled? Please put a (✓) Grade: 4 O 5 O 	
 4. Please put a (✓) who is completing this questionnaire: Other: 	Mother O Father O
 5. What language or languages are spoken at home? English O Spanish O Mandarin O Other(s) O (please specify)
6. In what language or languages does your child speak to c	others at home?

English O Spanish O Other (s) O (please specify)

7. In what country was your child born?

	Canada O	Outside Canada O	(please specify)	
If	your child was no	ot born in Canada, at wh	at age did your child n	nove to Canada?
8. Do ye	ou (mother) spea	k English? Yes O	No O	
H	ow much English	do you speak? a little	O so/so O	a lot O
9. Do ye	ou (father) speak	English? Yes O	No O	
H	ow much English	do you speak? a little	O so/so O	a lot O
10. If y	you (mother) we	ere not born in Canada	a, in what year did yo	u move to Canada?
11. If yo	ou (father) were	not born in Canada, in v	vhat year did you move	e to Canada?
12. How researc	v many other chil	dren live in the same ho	me as the child who is	participating in this
vv	hat are their ages	·		

13. Is your child generally healthy? Yes O No O

14. Does your child have a history of head injuries? Yes O No O

If your answer is yes, please explain.

15. Has your child's hearing been16. Do you have any concerns abo	ever tested? ut your child hearing?	Yes O No O Yes O No O	
17. Has your child repeated a grad If yes, tell us more	e? Yes C	No O	
18. Does your child have learning dif	fficulties? Yes O	No O	
If yes, tell us more			
19. Does your child receive speciaIf yes, please put a (✓) for	al help in learning at s Yes O each appropriate optic	chool? No O on.	
Extra support in class Withdrawal to learning centre ESL support Tutoring Speech therapy Other(S) O please specify_	At present O O O O O	In the past O O O O	
20. Does your child participate in literacy program, etc.)?O please specify	any extra-curricular a Yes O	ctivities (e.g. sports, private tuto No O	ring,
21. Please place a (✓) beside the h O Some school: O Completed high scho	ighest level of educati # of grades complete	on that you (mother) have attai d	ned.

- Completed high school diploma and a professional qualification not from a
- College or university
- Completed a college diploma
- Completed an undergraduate university degree
- Completed two or more university degrees
- O Other (please specify):

22. What is your (mother) occupation in Canada?_____

If you are a Canadian and were employed before immigrating to Canada, please indicate your occupation in your former country _____

- 23. Please place a (\checkmark) beside the highest level of education that you (<u>father</u>) have attained.
 - O Some school: _____ # of grades completed
 - Completed high school diploma
 - Completed high school diploma and a professional qualification not from a
 - O College or university
 - O Completed a college diploma
 - Completed an undergraduate university degree
 - Completed two or more university degrees
 - O Other (please specify): _____

24. What is your (father) occupation In Canada?

If you are a new Canadian and were employed before immigrating to Canada, please indicate your occupation in your former country _____

Thank you for completing the Family Demographic Questionnaire. We look forward to sharing the findings of the project with you.

PRE-INTERVIEW QUESTIONNAIRE

To be completed prior to either the one-on-one interview or focus group interview

1. Age 2. Sex (please circle): MALE FEMALE	3. Grade
4. First Language	
 4. Born in Canada? (please circle) YES NO 4a. If you were NOT born in Canada, how old were you when you came years old 	to Canada?
4b. If you were born in Canada , check one:	
I am a 1st generation Canadian (my parents were not born in Canad	a)
2nd generation Canadian (my parents were born in Canada)	
3rd generation Canadian (my grandparents were born in Canada)	
Other (please specify):	
5. Your cultural background -> Check ALL that apply:	
African-American / Black / African Origin	
Asian-American / Asian Origin / Pacific Islander	
Latino-a / Hispanic	
American Indian / Alaska Native / Aboriginal Canadian	
European Origin / White	
Bi-racial / Multi-racial	
Other (please specify):	

6. What five words or phrases would your closest friends use to describe your <u>personality</u>? WRITE YOUR FIVE PERSONALITY TRAITS BELOW:

1	_
2	
3	
4	
5	

	Daily	Few times a week	Once a week	Few times a month	Once a month	Less often	Never
Study for school							
Do homework							
Watch TV programs (English)							
Listen to the radio (English)							
Surf the Internet							
Read magazines (any language)							
Read books that are not for school purposes (any language)							
<u>Write</u> , but not for school purposes (ex. journal, blog, diary, letters, stories etc.)							
Email friends and family (any language)							
Text message friends and family (any language)							
Work or Volunteer (where:)							
Engage in other extra-curricular activities							
Another activity:							
Another activity:							
Another activity:							

6. How often do you participate in each of the following activities? PLEASE CHECK (\checkmark) ONE ANSWER PER ACTIVITY.

Great!

Thank you for taking the time to answer these questions.

B.3 Semi-structured Student Interview Guide

One-on-one Interview Guide

Please note that given the free response format, these questions will likely lead to additional questions to encourage elaboration of ideas, such as:

- Why?
- How often?
- Could you explain further?
- Would you give me an example of what you mean?
- Tell me more about that.
- How do you feel about that?
- What makes you feel that way?
- I'd like to know more about your thinking on that issue

Volunteering (warm-up question)

□ In general, how do you feel about volunteering for things (e.g. for community work, research projects, school clubs, etc.)?

VLP Program (Tutor, Group Dynamics, Preferences)

You [participated for _x__ session(s) in the VLP program/completed the VLP program].

- Did you attend every session? Why?
- Tell me about the things that you liked about the VLP.
- Tell me about the things that you did not like about the VLP.
- Why did you choose to complete/not complete the program?
- □ What do you think is the purpose of a program like the VLP?
- Tell me about your VLP tutor.
- □ I'd like to hear your thoughts about your tutoring group.
- What do you think you have learned from participating in the VLP?
- Do you have any suggestions about how we can make the VLP program better?
- □ What did you think about the length of the sessions? The size of your group?
- □ What are your thoughts on the reading materials that you used?
- □ Would you participate again in the VLP?

Peers and Family/Home Literacy

- Did any of your friends participate in the VLP? How did you feel about that?
- Do your friends help you with homework? Parents? Extended family help with homework, reading and writing?
- What kinds of things do your parents do to help you do well in school, stay motivated, etc.?
- Does your family read together at home or somewhere else (e.g. newspapers, magazines, religious materials, games, internet)?

U Who are most i	mportant people	in your life?
------------------	-----------------	---------------

Academic Self-Concept, Independent Literacy Engagement

- How do you feel about your performance in school, like your grades? Why?
- Do you enjoy reading in your spare time? Why?
- Do you ever practice the lessons you learned from VLP when you are reading or when you come across words you don't know?

School Factors

- □ Tell me about your school. What things do you like/not like about it?
- Tell me about your teachers. What things do you like/not like about them?
- Any teacher/mentor/friend/person who has made a lasting impact, in terms of helping with school, goal setting etc.?

Extra-curricular activities and responsibilities (if relevant)

- Tell me about your extracurricular activities. Do you think that participating in [insert activity] affected how well you did in the VLP? Does it affect how well you do at school?
- Did you participate in any academic summer program last summer? Tell me about it.

Language and Culture

- □ If English is not first language, ask: Do you speak or read [first language]?
- □ What language(s) do you usually speak with your friends?
- □ In what language are the TV, movie and radio programs that you prefer listening to or watching?
- How would you rate your English speaking and reading skills on a scale from 1 (poor) to 5 (excellent)? Why?

I'd like to ask about your cultural identity, meaning the culture(s) you feel you belong to and that you share your values and beliefs with.

When you or your family (e.g. parents/grandparents) come from a different country to live in Canada, often you are exposed to both cultures – Canadian culture and the other country's culture. Do you feel that you are more a part of Canadian culture, the other culture, or both? (Can you give me examples?)

Additional Questions Asked?

Focus group?

Thank you so much for taking the time to answer these questions. It really means a lot to us and will really help us understand more about how we can improve our reading programs.